# Catchment- and Event-Type Specific Synthetic Design Hydrographs

## Estimation, Characterization, Regionalization, and Uncertainty

**Dissertation**

zur

Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

**Universität Zürich**

von

**Manuela Irene Brunner**

von

Bettwil AG

Promotionskommission

**Prof. Dr. Jan Seibert (Vorsitz)**

**Prof. Dr. Anne-Catherine Favre (Vorsitz)**

**Prof. Dr. Reinhard Furrer**

**Zürich, 2018**

Communauté
UNIVERSITÉ Grenoble Alpes

Universität
Zürich[UZH]

# THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES**

**préparée dans le cadre d'une cotutelle *entre la Communauté Université Grenoble Alpes* et *l'Université de Zürich***

Spécialité : **Terre, univers, environnement/Océan, atmosphère, hydrologie**

Arrêté ministériel : le 6 janvier 2005 - 7 août 2006

Présentée par


## Manuela Irene BRUNNER


Thèse dirigée par **Anne-Catherine FAVRE** et **Jan SEIBERT**

préparée au sein de **l'Institut des Géosciences de l'Environnement (Grenoble) et de l'institut de Géographie (Zürich)**

dans **les Écoles Doctorales Terre Univers Environnement (Grenoble) et Géographie (Zürich)**

# Hydrogrammes synthétiques par bassin et types d'événements

## Estimation, caractérisation, régionalisation et incertitude

Thèse soutenue publiquement le **29.01.2018**,
devant le jury composé de :

**Prof. Veronica WEBSTER**
Professeure, Michigan Technological University, Rapportrice
**Prof. András BARDOSSY**
Professeur, Universität Stuttgart, Rapporteur
**Dr. Benoît HINGRAY**
CR CNRS, Institut des Géosciences de l'Environnement, Examinateur
**Prof. Reinhard FURRER**
Professeur, Universität Zürich, Co-encadrant de thèse
**Prof. Anne-Catherine FAVRE**
Professeure, Grenoble INP – Ense$^3$, Directrice de thèse
**Prof. Jan SEIBERT**
Professeur, Universität Zürich, Directeur de thèse
**Dr. Olivier OVERNEY**
Chef de division, Office fédéral de l'environnement, Invité

# Abstract

Design flood estimates are needed in hydraulic design for the construction of dams and retention basins and in flood management for drawing hazard maps or modeling inundation areas. Traditionally, such design floods have been expressed in terms of peak discharge estimated in a univariate flood frequency analysis. However, design or flood management tasks involving storage, in addition to peak discharge, also require information on hydrograph volume, duration, and shape. A bivariate flood frequency analysis allows the joint estimation of peak discharge and hydrograph volume and the consideration of their dependence. While such bivariate design quantiles describe the magnitude of a design flood, they lack information on its shape. An attractive way of modeling the whole shape of a design flood is to express a representative normalized hydrograph shape as a probability density function. The combination of such a probability density function with bivariate design quantiles allows for the construction of a synthetic design hydrograph for a certain return period which describes the magnitude of a flood along with its shape. Such synthetic design hydrographs have the potential to be a useful and simple tool in design flood estimation. However, the use of such hydrographs is faced by the following limitations. First, they rely on the definition of a bivariate return period which is not uniquely defined. Second, they usually describe the specific behavior of a catchment and do not express process variability represented by different flood types. Third, they are neither available for ungauged catchments nor are they usually provided together with an uncertainty estimate. To overcome these limitations, this thesis explores possibilities for the construction of synthetic design hydrographs in gauged and ungauged catchments and ways of representing process variability in design flood construction. Tools are proposed for both catchment- and flood-type specific design hydrograph construction and regionalization, and for the assessment of their uncertainty. The thesis shows that synthetic design hydrographs are a flexible tool allowing for the consideration of different flood or event types in design flood estimation. A comparison of different regionalization methods, including spatial, similarity, and proximity based approaches, showed that catchment-specific design hydrographs can be best regionalized to ungauged catchments using linear and nonlinear regression methods. It was further shown that event-type specific design hydrograph sets can be regionalized using a bivariate index flood approach. In such a setting, a functional representation of hydrograph shapes was found to be a useful tool for the delineation of regions with similar flood reactivities. An uncertainty assessment showed that the record length and the choice of the sampling strategy are major uncertainty sources in the construction of synthetic design hydrographs and that this uncertainty propagates through the regionalization process. This thesis highlights that an ensemble-based design flood approach allows for the consideration of different flood types and runoff processes. This is a step from flood frequency statistics to *flood frequency hydrology* which allows for the consideration of process variability and therefore better-informed decision making.

**Keywords:** Synthetic design hydrographs, bivariate flood frequency analysis, design flood estimation, regionalization, uncertainty assessment, clustering, homogeneous regions

# Zusammenfassung

Für hydraulische Problemstellungen und im Hochwassermanagement werden Bemessungshochwasser benötigt. Sie sind wichtig für die Planung von Dämmen und Retentionsbecken, um Gefahrenkarten zu erstellen und um Überflutungsflächen zu berechnen. Üblicherweise werden solche Bemessungshochwasser mit dem Wert der Hochwasserspitze ausgedrückt und mittels einer univariaten Analyse abgeschätzt. Für die Bemessung hydraulischer Strukturen oder Hochwassermanagementaufgaben werden aber zusätzlich zu der Hochwasserspitze auch Informationen über das Hochwasservolumen, die Dauer des Hochwassers und über die Form der Hochwasserganglinie benötigt. Eine bivariate Analyse erlaubt die gemeinsame Abschätzung von Hochwasserspitzen und –volumen unter Berücksichtigung deren gegenseitigen Abhängigkeit. Solche bivariaten Bemessungsquantile beschreiben die Grösse eines Ereignisses, nicht aber dessen Ganglinie. Die Ganglinie eines Hochwasserereignisses kann modelliert werden, indem ein repräsentativer, normalisierter Hydrograph durch eine Dichtefunktion ausgedrückt wird. Die Kombination einer Dichtefunktion mit bivariaten Bemessungsquantilen erlaubt die Konstruktion einer synthetischen Bemessungsganglinie mit einer bestimmten Wiederkehrperiode. Eine Bemessungsganglinie beschreibt die Grösse eines Hochwassers zusammen mit seiner Form. Solche Bemessungsganglinien erlauben eine einfache Abschätzung von Bemessungshochwassern, haben jedoch einige Limitationen. Erstens stützen sie sich auf eine Definition von einer bivariaten Wiederkehrperiode, welche nicht eindeutig definiert ist. Zweitens beschreiben sie das Verhalten eines Einzugsgebietes als Ganzes und erlauben es nicht die Variabilität von Prozessen in einem Gebiet zu berücksichtigen, welche durch verschiedene Hochwassertypen repräsentiert werden. Drittens sind sie weder für Gebiete ohne Abflussmessungen verfügbar noch werden sie üblicherweise zusammen mit Unsicherheitsangaben angegeben. In dieser Doktorarbeit wurden deshalb Möglichkeiten zur Konstruktion von synthetischen Bemessungsganglinien untersucht wobei sowohl gemessene als ungemessene Gebiete berücksichtigt wurden. Des Weiteren wurden Möglichkeiten untersucht um Prozessvariabilität in der Konstruktion von Bemessungshochwassern zu berücksichtigen. Es wurden Werkzeuge für die Konstruktion von gebiets- und ereignisspezifischen Bemessungsganglinien und deren Regionalisierung und Unsicherheitsbeurteilung vorgeschlagen. Die Arbeit hat aufgezeigt, dass synthetische Bemessungsganglinien ein flexibles Werkzeug sind, das die Berücksichtigung von verschiedenen Hochwasser- oder Ereignistypen erlaubt. Ein Vergleich von verschiedenen Regionalisierungsmethoden, der räumliche Methoden, sowie Methoden, die auf Ähnlichkeit oder Nähe basieren umfasste, hat gezeigt, dass gebietsspezifische Bemessungsganglinien am besten mit linearen oder nichtlinearen Regressionsmethoden regionalisiert werden können. Es wurde weiter aufgezeigt, dass ereignisspezifische Sets von Bemessungshochwassern mittels eines bivariaten Index Hochwasser Ansatzes regionalisiert werden können. In einem solchen Ansatz hat sich eine funktionale Repräsentation von Ganglinien als nützlich erwiesen, um Regionen mit ähnlichen Hochwasserreaktionen abzugrenzen. Eine Unsicherheitsanalyse hat gezeigt, dass die Länge der zur Verfügung stehenden Datenrei-

he und die Wahl des Hochwasserauswahlverfahrens die grössten Unsicherheitsquellen in der Konstruktion von Bemessungsganglinien sind und dass sich diese Unsicherheiten durch den Regionalisierungsprozess fortpflanzen. Ein Bemessungshochwasseransatz, der auf mehreren Ganglinien basiert, ermöglicht die Berücksichtigung von verschiedenen Hochwassertypen und Abflussprozessen. Dies ist ein Schritt von einer reinen Hochwasserhäufigkeitsstatistik in Richtung Hochwasserhäufigkeitshydrologie, welche den Einbezug von Prozessvariabilität erlaubt und somit die Entscheidungsfindung dank umfassender Information erleichtert.

**Schlüsselwörter:** Synthetische Bemessungsganglinien, bivariate Hochwasserhäufigkeitsanalyse, Regionalisierung, Unsicherheitsabschätzung, Clustering, homogene Regionen

# Résumé

L'estimation de crues de projet est requise pour le dimensionnement de barrages et de bassins de rétention, de même que pour la gestion des inondations lors de l'élaboration de cartes d'aléas ou lors de la modélisation et délimitation de plaines d'inondation. Généralement, les crues de projet sont définies par leur débit de pointe à partir d'une analyse fréquentielle univariée. Cependant, lorsque le dimensionnement d'ouvrages hydrauliques ou la gestion de crues nécessitent un stockage du volume ruisselé, il est également nécessaire de connaître les caractéristiques volume, durée et forme de l'hydrogramme de crue en plus de son débit maximum. Une analyse fréquentielle bivariée permet une estimation conjointe du débit de pointe et du volume de l'hydrogramme en tenant compte de leur corrélation. Bien qu'une telle approche permette la détermination du couple débit/volume de crue, il manque l'information relative à la forme de l'hydrogramme de crue. Une approche attrayante pour caractériser la forme de la crue de projet est de définir un hydrogramme représentatif normalisé par une densité de probabilité. La combinaison d'une densité de probabilité et des quantiles bivariés débit/volume permet la construction d'un hydrogramme synthétique de crue pour une période de retour donnée, qui modélise le pic d'une crue ainsi que sa forme. De tels hydrogrammes synthétiques sont potentiellement utiles et simples d'utilisation pour la détermination de crues de projet. Cependant, ils possèdent actuellement plusieurs limitations. Premièrement, ils reposent sur la définition d'une période de retour bivariée qui n'est pas univoque. Deuxièmement, ils décrivent en général le comportement spécifique d'un bassin versant en ne tenant pas compte de la variabilité des processus représentée par différents types de crues. Troisièmement, les hydrogrammes synthétiques ne sont pas disponibles pour les bassins versant non jaugés et une estimation de leurs incertitudes n'est pas calculée. Pour remédier à ces manquements, cette thèse propose des avenues pour la construction d'hydrogrammes synthétiques de projet pour les bassins versants jaugés et non jaugés, de même que pour la prise en compte de la diversité des types de crue. Des méthodes sont également développées pour la construction d'hydrogrammes synthétiques de crue spécifiques au bassin et aux événements ainsi que pour la régionalisation des hydrogrammes. Une estimation des diverses sources d'incertitude est également proposée. Ces travaux de recherche montrent que les hydrogrammes synthétiques de projet constituent une approche qui s'adapte bien à la représentation de différents types de crue ou d'événements dans un contexte de détermination de crues de projet. Une comparaison de différentes méthodes de régionalisation, notamment basées sur des approches spatiales, de similarité ou de voisinage, montre que les hydrogrammes synthétiques de projet spécifiques au bassin peuvent être régionalisés à des bassins non jaugés à l'aide de méthodes de régression linéaires et non linéaires. Il est également montré que les hydrogrammes de projet spécifiques aux événements peuvent être régionalisés à l'aide d'une approche d'indice de crue bivariée. Dans ce contexte, une représentation fonctionnelle de la forme des hydrogrammes constitue un moyen judicieux pour la délimitation de régions ayant un comportement hydrologique de crue similaire en termes de réactivité. Une analyse de l'incertitude

a montré que la longueur de la série de mesures et le choix de la stratégie d'échantillonnage constituent les principales sources d'incertitude dans la construction d'hydrogrammes synthétiques de projet, et que cette incertitude se propage dans le processus de régionalisation. Cette thèse démontre qu'une approche de crues de projet basée sur un ensemble de crues permet la prise en compte des différents types de crue et de divers processus. Ces travaux permettent de passer de l'analyse fréquentielle statistique de crues vers l'analyse fréquentielle hydrologique de crues permettant de prendre en compte les processus et conduisant à une prise de décision plus éclairée.

**Mots-clé:** hydrogrammes synthétiques de projet, analyse bivarée de crues, estimation de crues de projet, régionalisation, estimation de l'incertitude, classification, régions homogènes.

## Project embedding, papers, and author contributions

This doctoral thesis was conducted in the framework of a project on the estimation of flood volumes in Switzerland which was initiated by the Federal Office for the Environment (FOEN). The project aimed at providing an approach that enables practitioners to estimate design floods for a predefined return period in catchments with and without runoff observations. The approach should provide design floods that not only include information on the event magnitude in terms of peak discharge but also on the hydrograph volume and shape of the flood hydrograph. According to the FOEN, the main needs of the project were:

- Volumes of flooding to produce risk maps;
- Representative floods to compute sediment transport;
- Flood hydrographs to design several hydraulic works including retention basins or weirs for lake regulation;
- Design flood hydrographs to validate existing retention basin volumes in ungauged catchments;
- Representative durations of flows larger than a given threshold to define risk scenarios caused by erosion and to evaluate carried sediment amounts.

These needs were addressed by the research done during my PhD which deals with the estimation of synthetic design hydrographs in gauged and ungauged catchments, describing a flood event in terms of peak discharge, hydrograph volume, and its whole shape.

The work on this project resulted in six papers. Together with the summary text, they form this manuscript. The summary text first introduces the topics of synthetic design hydrograph construction, regionalization, and uncertainty. In a second chapter, the thesis objectives are defined. The third chapter gives an overview on the data used and on the main methods applied and developed in the thesis. In the fourth chapter, the most important results obtained in the six papers are summarized. Chapter five then discusses these results and shows limitations and perspectives of the approaches proposed. The summary text closes with a conclusion on the main contributions of this work. The summary text is followed by the six papers written in the context of this thesis. The papers are listed below and are referred to as Papers I to VI in the summary text. The author contributions are specified in the next section.

## List of papers

I. Brunner, M. I., J. Seibert, and A.-C. Favre (2016), Bivariate return periods and their importance for flood peak and volume estimation, *Wire's Water*, 3, 819–833, doi: 10.1002/wat2.1173

*Reprinted with permission of John Wiley and Sons*

II. Brunner, M. I., D. Viviroli, A. E. Sikorska, O. Vannier, A.-C. Favre, and J. Seibert (2017), Flood type specific construction of synthetic design hydrographs, *Water Resources Research*, 53, doi: 10.1002/2016WR019535

*Reprinted with permission of John Wiley and Sons*

III. Brunner, M. I., R. Furrer, A. E. Sikorska, D. Viviroli, J. Seibert, and A.-C. Favre (2017), Synthetic design hydrographs for ungauged catchments: A comparison of regionalization methods, *Stochastic Environmental Research and Risk Assessment*, doi: 10.1007/s00477-018-1523-3

*Reprinted with permission of Springer Nature*

IV. Brunner, M. I., A. E. Sikorska, R. Furrer, and A.-C. Favre (2017), Uncertainty assessment of synthetic design hydrographs for gauged and ungauged catchments, *Water Resources Research*, accepted for publication

V. Brunner, M. I., D. Viviroli, R. Furrer, J. Seibert, and A.-C. Favre (2017), Identification of flood reactivity regions via the functional clustering of hydrographs, *Water Resources Research*, accepted for publication

VI. Brunner, M. I., J. Seibert, and A.-C. Favre (2018), Representative sets of design hydrographs for ungauged catchments: A regional approach using probabilistic region memberships, *Advances in Water Resources*, 112, 235-244, doi: 10.1016/j.advwatres.2017.12.018

*Reprinted with permission of Elsevier Ltd.*

## Co-authorship

I had the lead in designing all the studies performed during my thesis. I implemented the statistical models and both analyzed and interpreted the results. This would not have been possible without the help of several people. The synthetic design hydrograph construction model was based on work done in a preliminary study of the project of the FOEN conducted by Anne-Catherine FAVRE, Paul MEYLAN, Jan SEIBERT, Anna SIKORSKA, Olivier VANNIER, and Daniel VIVIROLI. The runoff observations used in the five papers were requested from the FOEN and several cantons by Paul MEYLAN and Daniel VIVIROLI. The catchment characteristics data used in Papers III to VI were partly provided by Daniel VIVIROLI and partly computed by myself. The

code used for the peak-over-threshold sampling in the extraction of flood events was provided by Emmanuel PAQUET from Électricité de France (EDF). Preliminary tools for dependence analysis and synthetic design hydrograph construction were provided by Olivier VANNIER and an initial code for flood type classification was provided by Anna SIKORSKA. I wrote the first drafts of all six papers which were subsequently revised by the co-authors and language-checked by Tracy EWEN.

# CONTENTS

# Abbreviations and Variables

| | |
|---|---|
| AMH | Ali-Mikhail-Haq |
| $B$ | Baseflow |
| B-spline | Basis spline |
| $C$ | Copula |
| $D$ | Duration |
| DEM | Digital elevation model |
| $D_T$ | Design quantile duration |
| $E_{AR}$ | Absolute relative error |
| EDF | Électricité de France |
| $f(t)$ | Probability density function |
| $f(t_p)$ | Density at time to peak |
| FOEN | Federal Office for the Environment |
| $F_X$ | Marginal distribution of peak discharges |
| $F_Y$ | Marginal distribution of hydrograph volumes |
| $F_{XY}$ | Joint distribution of peak discharges and hydrograph volumes |
| FD | Functional data |
| $E_{\text{MAR}}$ | Mean absolute relative error |
| FF | Flash flood |
| GEV | Generalized extreme value distribution |
| GMF | Glaciermelt flood |
| GPD | Generalized Pareto distribution |
| $I_{BF}$ | Baseflow index |
| $k$ | Number of clusters |
| $\text{km}^2$ | Square kilometers |
| LRF | Long-rain flood |
| m.a.s.l. | Meters above sea level |
| $\mu$ | Location parameter |
| $\mu_i$ | Index flood |
| $\mu_t$ | Mean inter-arrival time |
| $N$ | Number of sites |
| PDF | Probability density function |
| POT | Peak-over-threshold |
| $q(F)$ | Regional growth curve |
| $Q_i(F)$ | Local quantile estimate |
| $Q_{\text{p}}$ | Peak discharge |
| $Q_T(t)$ | Synthetic design hydrograph |
| $Q_T$ | Design quantile peak discharge |
| $\mathbb{R}$ | Set of real numbers |
| RNH | Representative normalized hydrograph |
| RoSF | Rain-on-snow flood |
| SDH | Synthetic design hydrograph |
| SMF | Snowmelt flood |
| $\sigma$ | Scale parameter |
| SRF | Short-rain flood |
| $t$ | Time |
| $T$ | Return period |
| $t_{\text{p}}$ | Time to peak |
| $t_{\text{p05}}$ | Half-recession time |
| $\theta$ | Copula parameter |
| $u$ | $F_X(x)$ |
| UH | Unit hydrograph |
| $v$ | $F_Y(y)$ |
| $V$ | Volume |
| $V_T$ | Design quantile hydrograph volume |
| $x$ | Peak discharge threshold |
| $X$ | Peak discharge |
| $\xi$ | Shape parameter |
| $y$ | Hydrograph volume threshold |
| $Y$ | Hydrograph volume |

CHAPTER 1

# INTRODUCTION

August 2005. A flood triggered by widespread and heavy precipitation events causes damages of 3 Billion Swiss Francs in Switzerland. Such flood events are rare but can have severe consequences for society and economy. It is therefore in the interest of a society to protect settlements, infrastructure, and human life against such floods (*Bezzola and Hegg*, 2007). Flood hazard maps form the basis for flood risk assessments and show where the construction of buildings should be avoided to reduce damage potential. In addition to such planning measures, structural measures might be necessary to achieve the desired protection level. For the implementation of both types of measures, design flood estimates are required that describe the expected magnitude of a flood event with a certain recurrence interval or return period (*BWG*, 2003).

Design flood estimation has traditionally focused on peak discharge, which is often the main flood characteristic of interest, but provides only a coarse picture of a flood event (*Rosbjerg et al.*, 2013). Other hydrograph characteristics, such as volume, duration, and shape provide additional and potentially valuable information (*Mediero et al.*, 2010). Their consideration in design flood estimation is of particular interest if design floods are required for hydraulic design tasks involving storage, such as the construction of retention basins, or for flood management tasks, such as drawing hazard maps (*Pilgrim*, 1986) or modeling inundation areas. Different flood hydrograph shapes may cause differences in the costs of hydraulic structures and influence flood management strategies (*Yue et al.*, 2002). Design flood hydrographs provide a means to describe the physical properties of a flood with a specified recurrence interval or return period in terms of peak discharge, hydrograph volume, and shape (*Serinaldi and Grimaldi*, 2011). Herein the term *synthetic* design hydrographs (SDHs) is used to emphasize that design events summarize the flood behavior of a catchment in the form of an estimated, synthetic hydrograph, which has not

been actually observed. The estimation of SDHs is of interest for both gauged catchments, where runoff information is available, and ungauged catchments where such information is missing. Existing approaches for the construction of SDHs in gauged catchments and their regionalization to ungauged catchments are outlined in the sections below.

## 1.1 SDH construction

Design flood estimation methods can be grouped into statistical and deterministic methods or a combination of the two (*Rogger et al.*, 2012; *Smithers*, 2012). Statistical methods rely on flood frequency analyses. They fit a mathematical probability distribution to the design variable of interest, which is often peak discharge, and extrapolate the tails of this distribution to low exceedance probabilities (*Klemes*, 1993). They require long observed flood records to avoid unreliable estimates (*Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012). On the contrary, deterministic methods are event based. They define a design hydrograph through the transformation of rainfall into runoff. The rainfall input is a design rainfall hyetograph with an assigned return period deduced from an intensity-duration-frequency curve (*Grimaldi et al.*, 2012a; *Soczynska et al.*, 1997). This approach allows the consideration of catchment processes. However, it is based on three critical assumptions (*Pilgrim and Cordery*, 1993): the choice of the design rainfall hyetograph, the equality between the rainfall and discharge return periods (*Viglione et al.*, 2009), and the choice of initial soil moisture conditions (*Camici et al.*, 2011). Combined approaches stochastically generate long rainfall and temperature time series and put these into a continuous rainfall-runoff model to simulate long runoff time series, which are then used in flood frequency analysis (*Grimaldi et al.*, 2012b). Such combined approaches avoid making the critical assumptions of deterministic approaches. However, they require, on the one hand, the choice of a suitable stochastic model and on the other hand, the calibration of a continuous rainfall-runoff model, which is for a practitioner, more time consuming than classical flood frequency analysis and therefore unlikely to be used for the design of minor works (*Boughton and Droop*, 2003). Because of the critical assumptions behind the deterministic approach and the time-constraint of practitioners, this thesis focuses on statistical methods that can be easily applied by practitioners and produce reproducible results. Contrary to classical statistical approaches, the methods developed in this thesis focus on a bivariate frequency analysis, which allows the joint consideration of peak discharges and hydrograph volumes. In addition, they consider the whole hydrograph shape without reducing it to a few hydrograph characteristics. Hydrograph shapes have traditionally been modeled by unit hydrographs (UHs). Methods for UH derivation comprise traditional unit hydrographs, synthetic unit hydrographs, typical hydrographs, and statistical methods (*Yue et al.*, 2002). Statistical methods are based on probability density functions (PDFs) fitted to observed hydrographs and have been found to be more suitable to derive unit hydrographs than traditional methods because they are flexible,

can take various shapes, and the area under the curve is guaranteed to be equal to one. These characteristics make PDFs a suitable basis for design flood hydrographs (*Bhunya et al.*, 2007). A dimensionless UH can be combined with the design variables (peak discharge and hydrograph volume) estimated in a bivariate frequency analysis to derive an actual design flood hydrograph with a specified return period (*Serinaldi and Grimaldi*, 2011).

Existing design flood estimation methods, independent of the estimation method chosen, focus on catchment-specific design floods which summarize the flood behavior of a catchment. However, floods can be triggered by different processes represented by different flood types (*Merz and Blöschl*, 2003). These yield important information on the flood behavior of a catchment. Classical SDH construction methods typically do not allow for the explicit consideration of different flood types such as flash floods, short-rain floods, long-rain floods, or rain-on-snow floods (*Diezig and Weingartner*, 2007; *Merz and Blöschl*, 2003; *Sikorska et al.*, 2015). However, including such causal information (*Merz and Blöschl*, 2008a) in flood frequency analysis and SDH construction could provide more reliable flood estimates.

## 1.2 Regionalization

In contrast to gauged catchments, ungauged catchments lack runoff observations which could be used in flood frequency analysis or SDH construction. Therefore, alternative data such as catchment characteristics and climate or spatial data are usually used to transfer or regionalize flood estimates from gauged to ungauged catchments (*Blöschl et al.*, 2013). Regionalization methods comprise 1) methods establishing a relation between catchment characteristics and model parameters, 2) approaches based on spatial proximity, and 3) methods delineating regions of catchments with a similar flood behavior (*Steinschneider et al.*, 2014). A range of different data types has previously been regionalized to ungauged catchments including unit hydrograph parameters (*Tung et al.*, 1997) and flood quantiles (*Merz and Blöschl*, 2004; *Ouarda et al.*, 2001; *Skoien et al.*, 2006) among other hydrological variables. However, very few studies deal simultaneously with both multivariate and regional aspects. *Requena et al.* (2016) proposed an approach for regionally estimating design quantiles via a bivariate regional index flood approach. The index flood approach was originally proposed by *Dalrymple* (1960) to predict univariate design quantiles using a pool of data from a region of similar catchments. It consists of two main steps. In a first step, regions with a similar flood behavior are delineated. In a second step, the data within these similar regions are used for regional flood frequency analysis. It assumes that frequency distributions at different sites within a region are identical apart from a scale factor. It describes a local quantile estimate $Q_i(F)$ as the product of an index flood ($\mu_i$) and a regional growth curve ($q(F)$) estimated based on the data at $N$ sites so that:

$$Q_i(F) = \mu_i q(F) \qquad i = 1,\ldots,N. \tag{1.1}$$

The index flood can be any location parameter of the at-site distribution but is often taken to be its mean. The regional growth curve is a dimensionless quantile function computed based on dimensionless regional data, which are obtained by dividing the observed flood event data by the index flood. The classical index flood procedure focuses on peak discharges. *Requena et al.* (2016) therefore proposed an approach for a multivariate regional index flood approach to allow for the consideration of more than one design hydrograph characteristic, e.g. peak discharge and hydrograph volume. While such a bivariate regional approach allows the joint consideration of peak discharges and hydrograph volumes, neither hydrograph shape, nor process variability can be considered. To the best of my knowledge, no method has so far been proposed for the regionalization of SDHs that represent peak discharge and flood volume together with the whole hydrograph shape. Furthermore, no methodology has so far been proposed for the regional construction of event-type specific sets of SDHs in ungauged catchments.

## 1.3 Uncertainty

Flood estimation entails various sources of uncertainty comprising measurement errors, various assumptions, sample selection, the choice of a suitable distribution function, the choice of a parameter estimation method, and sampling uncertainty (*Merz and Blöschl*, 2005). Such uncertainty sources interact and together determine the total uncertainty of flood estimates (*Beven and Hall*, 2014; *Merz and Blöschl*, 2008a). Uncertainty analysis allows the identification of uncertain parameters (*Tung and Yen*, 2005), a quantitative assessment of model reliability (*Merz and Blöschl*, 2005; *Montanari and Koutsoyiannis*, 2012), and indicates where potential improvements in the method could have the greatest impact (*Cullen and Frey*, 1999; *Hall and Solomatine*, 2008; *Sikorska et al.*, 2012). Still, uncertainty assessments are not routinely conducted in flood frequency analyses (*Pappenberger et al.*, 2006).

CHAPTER 2

# THESIS OBJECTIVES

There is a need for a simple design hydrograph estimation method that can be easily applied by engineers with little hydrological expertise for the design of minor works, such as bridges and culverts, farm dam spillways, and urban drainage systems, where a certain risk of failure is acceptable (*Pilgrim and Cordery*, 1993). Such a method should be applicable both in gauged and ungauged catchments and represent peak discharge and hydrograph volume together with the whole hydrograph shape (*Grimaldi et al.*, 2012a,b; *Mediero et al.*, 2010). Furthermore, such a method should quantify the uncertainty of an SDH (*Beven et al.*, 2010). For these reasons, this PhD thesis developed an SDH construction method for gauged catchments solely based on runoff observations. The SDHs characterize the design flood not only in terms of peak discharge but also hydrograph volume and shape. In a second step, this method was regionalized to ungauged catchments to enable the estimation of SDHs based on catchment characteristics in the absence of runoff observations. In a third step, a simulation framework was set up to assess the uncertainty of SDHs constructed in gauged catchments and regionalized to ungauged catchments. The method was not limited to the construction of catchment-specific SDHs but was extended to an ensemble-based approach that allows for the consideration of different process types occurring in a catchment.

More specifically, the PhD thesis addressed the following research questions:

**Question 1** How can the dependence between peak discharges and hydrograph volumes be jointly considered in design hydrograph construction together with a realistic representation of the hydrograph shape?

**Question 2** How can process types be considered in design hydrograph construction?

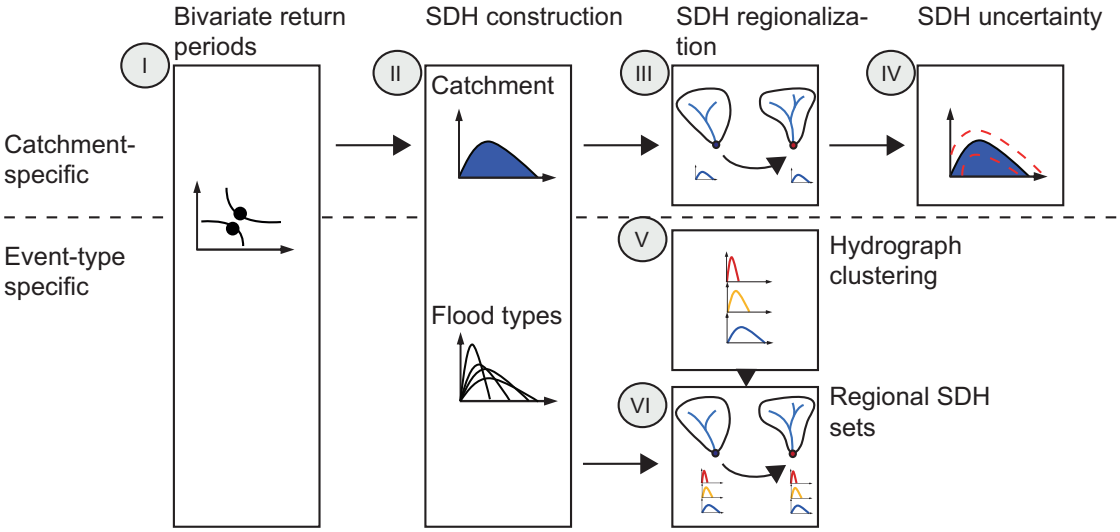



Figure 2.1: Overview of the six papers written in the context of this PhD thesis and links between them.

**Question 3** What is the most suitable method to regionalize SDHs from gauged to ungauged catchments and how can process types be considered in the regionalization to ungauged catchments?

**Question 4** What are the most important uncertainty sources in design hydrograph construction and what is the uncertainty of SDH estimates in gauged and ungauged catchments?

These research questions were addressed in six scientific papers which build up on each other (Figure 2.1).

**Paper I** contributes to answering Question 1 and is essential for answering all remaining questions. It deals with the question of how to define a return period in a bivariate context which is fundamental in the design of SDHs representing the two design variables peak discharge and hydrograph volume. It reviews tools that are used to define return periods in a bivariate context and illustrates the effect of different return period definitions on bivariate design quantile estimates in a case study.

**Paper II** deals with Questions 1 and 2. It proposes a simple probabilistic SDH construction framework that is based on observed runoff data only and avoids the use of any rainfall-runoff model. The constructed SDHs describe design floods in terms of peak discharge and hydrograph volume via a bivariate frequency analysis and the hydrograph shape via a probability density function. The catchment-specific construction approach is extended to a flood-type specific approach where SDHs are constructed for different flood types such as long-rain, short-rain, rain-on-snow, and flash floods.

**Paper III** deals with Question 3 and tries to identify a suitable method for the regionalization of SDHs to ungauged catchments. Therefore, it compares various methods for the regionalization of catchment-specific SDHs to ungauged catchments comprising methods establishing a relation-

ship between SDH parameters and catchment characteristics, spatial methods, and methods based on the delineation of homogeneous regions.

**Paper IV** deals with Question 4. It proposes an uncertainty assessment framework with three levels of complexity which allows for the identification of the most important uncertainty sources in the construction and regionalization of SDHs. Furthermore, it enables the quantification of the total uncertainty of constructed and regionalized SDHs and propagates construction uncertainty through regionalization.

**Paper V** deals with Questions 2 and 3. It proposes a clustering approach based on functional data to identify sets of representative hydrograph shapes within a catchment and regions of similar flood reactivity on a regional scale. The delineation of such reactivity regions is a first step towards a regional SDH construction approach that allows for the construction of event-type specific SDHs in ungauged catchments.

**Paper VI** deals with Question 3. It proposes a method for the regionalization of event-type specific SDH sets to ungauged catchments. It uses the idea of the delineation of flood reactivity regions proposed in Paper V in a regional index flood approach which is extended to a regional SDH construction approach.

The six papers form a logical sequence and are closely linked. Bivariate return periods as discussed in Paper I are fundamental for the construction of SDHs which is discussed in Paper II for both a catchment-specific and a flood-type specific framework. Paper III focuses on the regionalization of catchment-specific SDHs. In contrast, Papers V and VI deal with the regionalization of event-type specific SDHs. In a first step, Paper V deals with the delineation of regions with similar representative hydrograph sets and therefore similar reactivity. In a second step, Paper VI uses these regions in a regional SDH construction approach which is based on a bivariate index flood approach. Paper IV looks at the uncertainty of SDHs focusing on catchment-specific SDHs. Uncertainty due to process variability is addressed in papers V and VI by the construction of event-type specific SDHs. All six papers focus on a return period of $T = 100$ years since this return period is often used as a reference in national guidelines to define protection goals for settlements in Switzerland (*Camezind-Wildi*, 2005). The methods proposed or applied, however, are applicable for any return period of interest considering the limits posed by the limited data availability (*Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012).

CHAPTER 3

# MATERIAL AND METHODS

This chapter first introduces the dataset used to perform the analyses conducted within this thesis. An overview is then given on the methods used to construct SDHs in gauged catchments, to regionalize SDHs to ungauged catchments, and to assess SDH uncertainty.

## 3.1 Study catchments

The analyses conducted in the context of this thesis rely on observed runoff time series and catchment characteristics of 163 study catchments in Switzerland.

### 3.1.1 Runoff observations

The dataset comprises catchments (Figure 3.1) with a wide range of catchment characteristics and flood behaviors. The selected catchments have hourly flow series of at least 20 years in duration ranging up to 53 years. The catchments' runoff is neither altered by regulated lakes upstream or inland canals nor by urbanized areas. The catchments are small to medium-size (6 to 1800 $km^2$), situated at mean elevations between 400 and 2600 m.a.s.l., and have either no or only a few areas with glaciers. Papers III to VI used the whole set of catchments while Papers I and II used only a subset of catchments.

The basis for each analysis was samples of flood events extracted from the runoff time series of the study catchments. Flood events were sampled using a peak-over-threshold approach based on the procedure proposed by *Lang et al.* (1999). The threshold for the peak discharges was chosen iteratively to fulfill a target condition of four events per year on average which is a trade-off between maximizing the information content in the sample and keeping the assumption of independence between events. For each of these events, sampled according to the flood peaks,

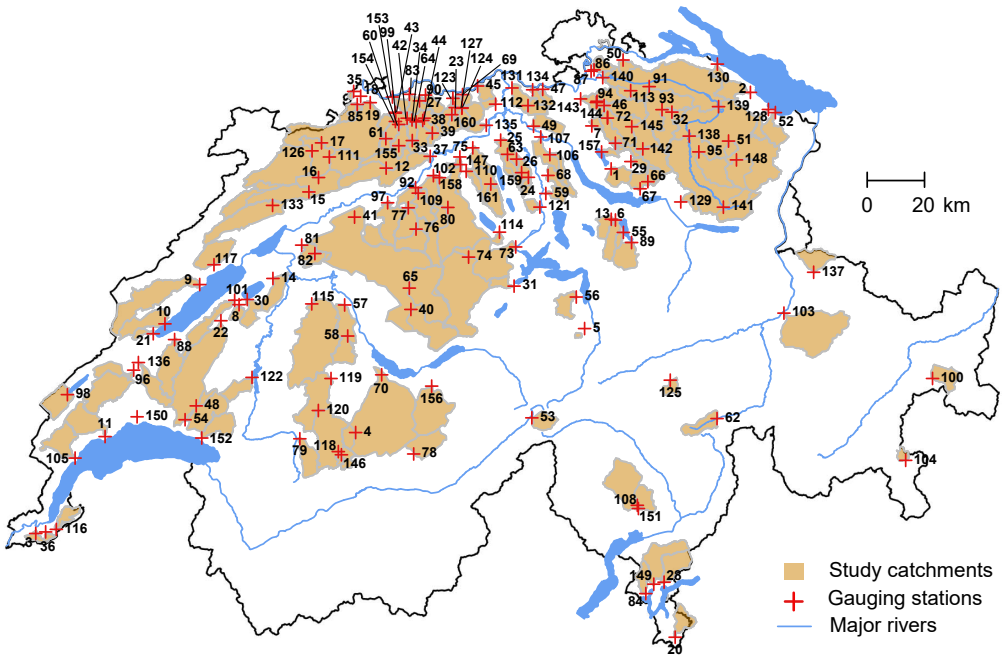



Figure 3.1: Map of the 163 Swiss study catchments used in this PhD thesis. The gauging stations are indicated as red crosses and labeled with a catchment ID (see Table 1 in the Appendix List of Catchments).

the flood volume and the whole hydrograph were determined over a fixed event window of 72 hours.

### 3.1.2 Catchment characteristics

A set of 54 catchment characteristics was used for the regionalization of SDHs to ungauged catchments. These characteristics comprised characteristics related to the physiography, land-use, soil, hydrogeology, geology, climate, and population of a catchment. Most attributes were computed using the PREVAH pre-processing tool WINHRU (*Viviroli et al.*, 2009). WINHRU derived physiographical characteristics from the digital elevation model (DEM), land-use related attributes from digital maps of land-use (*Bundesamt für Statistik*, 2003), soil related characteristics from digital maps of land surface characteristics (*Eidgenössische Forschungsanstalt für Wald Schnee und Landschaft (WSL)*, 1999), hydrogeology related characteristics from a map that focuses on groundwater resources (*Bitterli et al.*, 2007), and geology related attributes from the Swiss geotechnical map (*Bundesamt für Statistik*, 2003). The climatological characteristics were computed using tables from the Hydrological Atlas of Switzerland (*Jensen et al.*, 1997) and gridded meteorological data provided by MeteoSwiss (*MeteoSwiss*, 2013). The population density was computed using a population map for Switzerland (*Bundesamt für Statistik*, 2003). A complete list of the 54 catchment characteristics is provided in Table 2 of Paper III. Paper VI used a subset of these available catchment characteristics for establishing a classification rule

allowing the assignment of an ungauged catchment to one of the reactivity regions. This subset consisted of catchment area, network density, Y-coordinate, soil topographic index, percentage area of karstic rocks, sunshine duration, and vapor pressure.

## 3.2 Bivariate return periods

The return period contains information about the non-exceedance probability of an event which has traditionally been described in terms of its peak discharge in a univariate frequency analysis. A bivariate analysis, however, is more appropriate if two dependent variables, such as peak discharge and hydrograph volume, play a significant role in the ruling behavior of a flood (*Salvadori et al.*, 2014). While the return period is uniquely defined in a univariate setting, this is not the case in the bivariate setting (*Gräler et al.*, 2013; *Salvadori et al.*, 2011; *Yue and Rasmussen*, 2002). The return period used to describe bivariate events can be determined by three types of approaches: 1) use of the conditional probability, 2) use of the joint probability, and 3) use of the Kendall's distribution or survival function. The choice of one of these approaches depends on the problem at hand (*Serinaldi*, 2015). The conditional probability is used if one variable (often peak discharge) is considered to be more important than the other variable of interest (e.g. hydrograph volume). On the contrary, the joint probability is used if both variables are considered to be equally important for the design problem at hand. All these bivariate approaches have in common that, on the contrary to the univariate approach, there is no unique solution for design variables with a specific return period $T$. Instead, various combinations of peak discharges and hydrograph volumes share the same return period and lie on an isoline. For practical reasons, one event on the isoline is often chosen and called the design event. Usually, this is the most-likely design realization on the isoline, i.e., the point with the largest probability density along the isoline (*Salvadori et al.*, 2011). A more detailed description of possible return period definitions and of approaches to choose one design variable pair is provided in Paper I.

This thesis considers both peak discharge and hydrograph volume to be equally important for the design problem at hand and therefore focuses on the joint OR return period which uses the probability that either peak discharge or hydrograph volume (or both) exceeds a given threshold. It is defined as:

$$T(x,y) = \frac{\mu_t}{\Pr[X > x \vee Y > y]} = \frac{\mu_t}{1 - F_X(x) - F_Y(y) + F_{XY}(x,y)} = \frac{\mu_t}{1 - C(u,v)}, \tag{3.1}$$

where $X$ and $Y$ are random variables, $C$ is a copula, $x$ and $y$ are given thresholds, $\mu_t$ is the inter-arrival time between two successive events $u = F_X(x)$ and $v = F_Y(y)$, and $F_X$, $F_Y$, and $F_{XY}$ are the marginal and joint distribution functions of the random variables respectively (*Salvadori*, 2004; *Salvadori and De Michele*, 2004). Throughout this thesis, one single design variable pair is obtained from the isoline by selecting the point with the largest joint probability density as follows:

$$(u,v) = \underset{C(u,v)=t}{\operatorname{argmax}} f_{XY}\{F_X^{-1}(u), F_Y^{-1}(v)\}. \tag{3.2}$$

### 3.2.1 Marginal distributions

The marginal distributions of peak discharges and hydrograph volumes are linked to the flood sampling strategy. The flood samples used in this thesis were obtained by a peak-over-threshold (POT) approach which samples those events with a peak discharge higher than a predefined threshold (*Lang et al.*, 1999). POT samples, according to extreme value theory (*Coles*, 2001), follow a generalized Pareto distribution (GPD) which is defined as:

$$F_X(x) = 1 - \left\{1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}} \qquad \xi \neq 0, \tag{3.3}$$

with a location parameter $\mu$ in $\mathbb{R}$, a scale parameter $\sigma > 0$, and a shape parameter $\xi$ in $\mathbb{R}$.

The corresponding hydrograph volumes were defined over a time window of 72 hours starting 24 hours before peak discharge. Their marginal distribution can be described by a generalized extreme value (GEV) distribution as:

$$F_Y(y) = \exp\left[-\left\{1 + \xi\left(\frac{y-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}}\right] \qquad \xi \neq 0, \tag{3.4}$$

with domain $1 + \xi\left(\frac{y-\mu}{\sigma}\right) > 0$ for $\xi \neq 0$ and with a location parameter $\mu$ in $\mathbb{R}$, a scale parameter $\sigma > 0$, and a shape parameter $\xi$ in $\mathbb{R}$.

POT series have the advantage of including all relevant flood events, which is not guaranteed when working with annual maxima series (*Lang et al.*, 1999). However, the independence of the events needs to be ensured. In this thesis, a minimum time difference of 72 hours between two successive events was prescribed to ensure their independence.

### 3.2.2 Copulas

The joint distribution of peak discharges and hydrograph volumes needs to reflect their dependence. Contrary to classical bivariate distributions, copula models allow the modeling of two dependent variables with different marginal distributions (*Genest and Favre*, 2007). The copula approach has its origin in the representation theorem of *Sklar* (1959) which states that the joint cumulative distribution function $F_{XY}(x,y)$ of any pair of continuous random variables $(X,Y)$ can be written as:

$$F_{XY}(x,y) = C\{F_X(x), F_Y(y)\}, x, y \in \mathbb{R}, \tag{3.5}$$

where $F_X(x)$ and $F_Y(y)$ are the marginal distributions and $C : [0,1]^2$ is the copula. $C$ is unique if the marginals are continuous. One of the main advantages of the copula approach is that the

selection of an appropriate copula for modeling the dependence between $X$ and $Y$ can proceed independently from the choice of their marginal distributions (*Genest and Favre*, 2007).

The copula models considered in the context of this thesis were five copulas of the Archimedean copula family (Ali-Mikhail-Haq (AMH), Clayton, Frank, Gumbel (also belonging to the extreme-value copula family), and Joe), two copula models of the elliptical copula family (Normal and Student-$t$ copula), and the independence copula (*Joe*, 2014). They were fitted using maximum pseudo-likelihood estimation to the pseudo-observations, which are deduced from the ranks of the observations. After the fitting, the copulas were tested using both graphical approaches and a goodness-of-fit test based on the Cramér-von-Mises statistic (*Genest and Favre*, 2007). A $p$-value for the Cramér-von-Mises statistic of each copula was estimated using a statistical bootstrap procedure (*Genest et al.*, 2009). The Joe copula was found to be the most suitable model to represent the dependence between peak discharges and hydrograph volumes within a single catchment. It is expressed as:

$$C(u,v) = 1 - \left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right]^{\frac{1}{\theta}}, \tag{3.6}$$

where $\theta$ is the copula parameter, $u = F_X(x)$ and $v = F_Y(y)$ are uniformly distributed between 0 and 1, and their dependence is modeled by the copula $C$.

The Joe copula is very flexible and can represent the bivariate distributions of the flood samples in most catchments even when they are divided by flood types. In addition, it is able to model the upper tail dependence (*Heffernan*, 2000) present in the data as described by the upper tail dependence coefficient (*Poulin et al.*, 2007).

In a regional setting (see Paper VI), when pooling data from several catchments, the Student-$t$ copula was found to be more suitable to model the dependence in the data than the Joe copula. It showed the lowest Cramér-von-Mises test statistic among all the copulas tested and similar to the Joe copula was able to model the tail dependence in the data (*Frahm et al.*, 2003).

## 3.3 SDH construction

Catchment-specific and flood-type specific SDHs are constructed using observed hourly runoff time series. The construction approach models the magnitude of the design flood using a bivariate frequency analysis and the shape of the hydrograph via a probability density function. The method can be either applied to construct a catchment-specific SDH without any differentiation between flood types or to construct flood-type specific SDHs. Flood-type specific SDHs can be constructed for flash floods (FF), short-rain floods (SRF), long-rain floods (LRF), or rain-on-snow floods (RoSF), but not for snowmelt floods (SMF) and glaciermelt floods (GMF). The eleven steps involved in the procedure are illustrated in Figure 3.2, listed in the paragraphs below, and described in detail in Paper II. A more detailed description of the catchment-specific construction procedure can be found in *Brunner et al.* (2016).

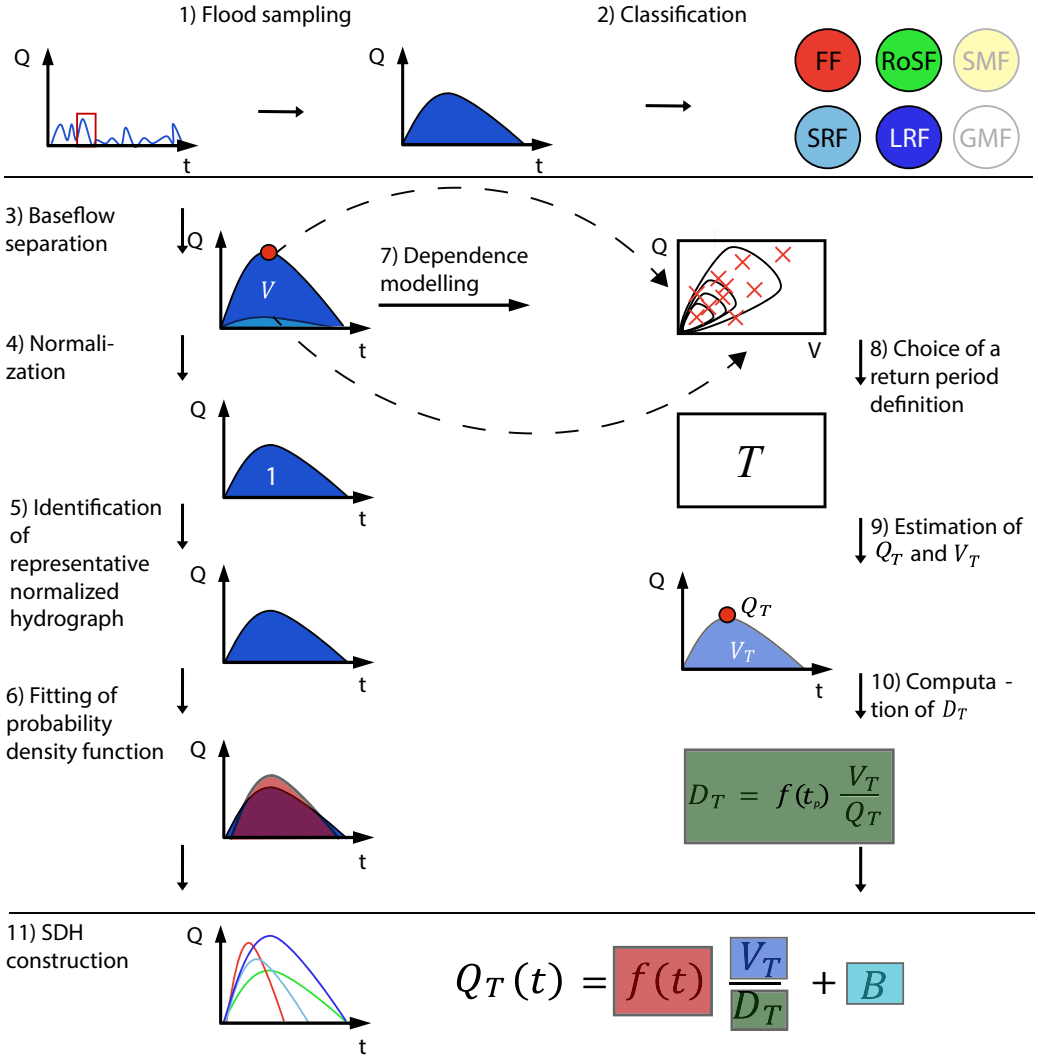



Figure 3.2: Method for the construction of synthetic design hydrographs (SDHs). The method can either be applied to all types of flood events identified in a catchment or it can be used to construct flood-type specific SDHs. The approach consists of eleven steps. The SDH ($Q_T(t)$) can be expressed by a probability density function ($f(t)$) times the mean discharge ($V_T/D_T$) plus the baseflow ($B$).

1. **Flood sampling:** Flood sampling using a POT approach which samples four events per year on average (see also Section 3.2.1).

2. **Classification:** Attribution of each flood event to one out of six flood types according to its triggering mechanism using a flood classification scheme adjusted from *Merz and Blöschl* (2003) and *Sikorska et al.* (2015). The flood types comprise flash floods (FFs) caused by short but very intense rainfall, short-rain floods (SRFs) caused by rainfall lasting no longer than one day, long-rain floods (LRFs) caused by rainfall lasting several days, rain-on-snow floods (RoSFs) triggered by rainfall falling on snow, snowmelt floods (SMFs) caused by a temperature-induced melting of snow cover, and glaciermelt floods (GMFs) caused by a temperature-induced melting of glacier cover. The classification scheme assigns a flood event to one of the flood types based on eight indices: timing of the flood, precipitation amount, precipitation duration, precipitation intensity, glacier cover, snow cover, snowmelt, and catchment wetness (*Sikorska et al.*, 2015). The focus of this thesis was on FFs, SRFs, LRFs, and RoSFs because the number of SMFs and GMFs was very low and the latter are less relevant for flood management because of their rather low peak discharges. The classification step can be skipped if catchment-specific SDHs instead of flood-type specific SDHs are of interest.

3. **Baseflow separation:** Separation of baseflow from the quick flow component of the event hydrographs to enable their statistical analysis (*Yue et al.*, 2002). A recursive digital filter was applied (*Eckhardt*, 2005) to separate quick flow from baseflow whose two parameters were estimated for each catchment individually.

4. **Normalization:** Normalization of the quick flow component of the hydrographs by dividing the base width by its duration $D$ and the ordinate by the mean runoff ($V/D$).

5. **Identification of representative normalized hydrograph:** Identification of a representative normalized hydrograph (RNH) as the median normalized hydrograph of the event set under consideration (whole dataset or flood-type specific datasets). The median hydrograph was defined using a notion of depth for functional data (*Ramsay and Silverman*, 2002). The concept of data depth can be used to define the centrality of a hydrograph within a group of hydrographs and to define their ranks. This enables the computation of robust estimators of the median among curves. Among different definitions of data depth, the $h$-mode depth (*Cuevas et al.*, 2007) was chosen and used to determine the median hydrograph within a set of hydrographs.

6. **Fitting of probability density function:** Fitting of a probability density function (PDF) to the RNH. PDFs can take various shapes and the area under their curves is equal to one as is the one under the RNH. The parameters of the PDF can be derived via an analytical expression using the time to peak, peak discharge, and the time base of the

RNH (*Nadarajah*, 2007; *Rai et al.*, 2009). A comparison of eight commonly used PDFs (beta, gamma, Fréchet, inverse gamma, logistic, lognormal, normal, and Weibull) has shown that the lognormal is the most suitable PDF to model the shape of the hydrographs. The dimensionless shape of the design hydrograph can be upscaled to an SDH using design variable quantiles.

7. **Dependence modeling:** Estimation of the parameters of the marginal distributions of peak discharges (GPD: Equation 3.3) and hydrograph volumes (GEV: Equation 3.4). Estimation of the parameter of the Joe copula. Modeling of the dependence between peak discharges and hydrograph volumes using the Joe copula model (Equation 3.6).

8. **Choice of return period definition:** Choice of a return period definition according to the problem at hand (*Serinaldi*, 2015) (see Section 3.2). In this thesis the joint OR return period was used assuming that peak discharge and hydrograph volume were equally important.

9. **Estimation of $Q_T$ and $V_T$:** Estimation of the design variable quantiles peak discharge ($Q_T$) and hydrograph volume ($V_T$) with the defined return period. The quantiles are obtained by inverting their marginal distributions $F_X$ and $F_Y$:

$$Q_{\mathrm{T}} = F_X^{-1}(u) \tag{3.7}$$

and

$$V_T = F_Y^{-1}(v). \tag{3.8}$$

10. **Computation of $D_T$:** Computation of the design duration ($D_T$) from the design estimates of $Q_T$ and $V_T$ as $D_T = f(t_{\mathrm{p}})V_T/Q_T$, where $f(t_{\mathrm{p}})$ is the lognormal density at the time to peak $t_{\mathrm{p}}$.

11. **SDH construction:** Construction of the SDH using the lognormal distribution fitted to the RNH and the estimates for $V_T$ and $D_T$ using:

$$Q_T(t) = f(t)V_T/D_T + B, \tag{3.9}$$

where $B$ is the baseflow to be added to the direct flow component. The baseflow is estimated by a mean event baseflow index computed for each catchment.

The construction of an SDH requires knowledge of ten parameters: event baseflow index, location and scale parameter of the lognormal distribution, a location, scale, and shape parameter for the two marginal distributions of the peak discharges and the hydrograph volumes as well as the parameter $\theta$ of the Joe copula. A more detailed description of the statistics of these parameters, their correlation, and distributions is provided in Paper III. The approach described above can be applied to the whole sample of flood events in a catchment or to a subsample of flood events representing a particular flood type. The event-type specific construction is an ensemble based approach that represents the flood behavior of a catchment described by different process types.

## 3.4 SDH regionalization

Design flood estimates are not only required in catchments where runoff observations are available but also in ungauged catchments where such observations are missing. Therefore, suitable approaches were sought that allowed for the transfer of SDHs to ungauged catchments. In a first step, catchment-specific SDHs were regionalized to ungauged catchments. In a second step, event-type specific SDH sets were transferred to ungauged catchments. The approaches used for the regionalization of catchment-specific SDHs is summarized in this paragraph and described in more detail in Paper III. The regionalization of event-type specific SDHs is addressed in Section 3.7 and described in detail in Paper VI. Various regionalization approaches can be used to regionalize the ten parameters of a catchment-specific SDH (see Section 3.3) to ungauged catchments. In this thesis, approaches from three categories were tested for their suitability to regionalize SDH parameters: 1) methods based on the relation between catchment characteristics and model parameters, 2) approaches based on spatial proximity, and 3) methods based on homogeneous regions (Figure 3.3). Approaches belonging to Category 1 comprise linear and nonlinear regression models. Methods belonging to Category 2 comprise spatial methods such as $k$-nearest neighbors, inverse distance weighting, and several kriging approaches. Category 3 delineates homogeneous regions, assigns an ungauged catchment to one of the regions based on its catchment characteristics, and estimates its SDH parameters based on the SDH parameters of the catchments belonging to the corresponding region. Since nonlinear regression models have not been frequently used in hydrological applications, they are shortly described in the next paragraph. More information on the other regionalization methods can be found in Paper III and the references cited therein.

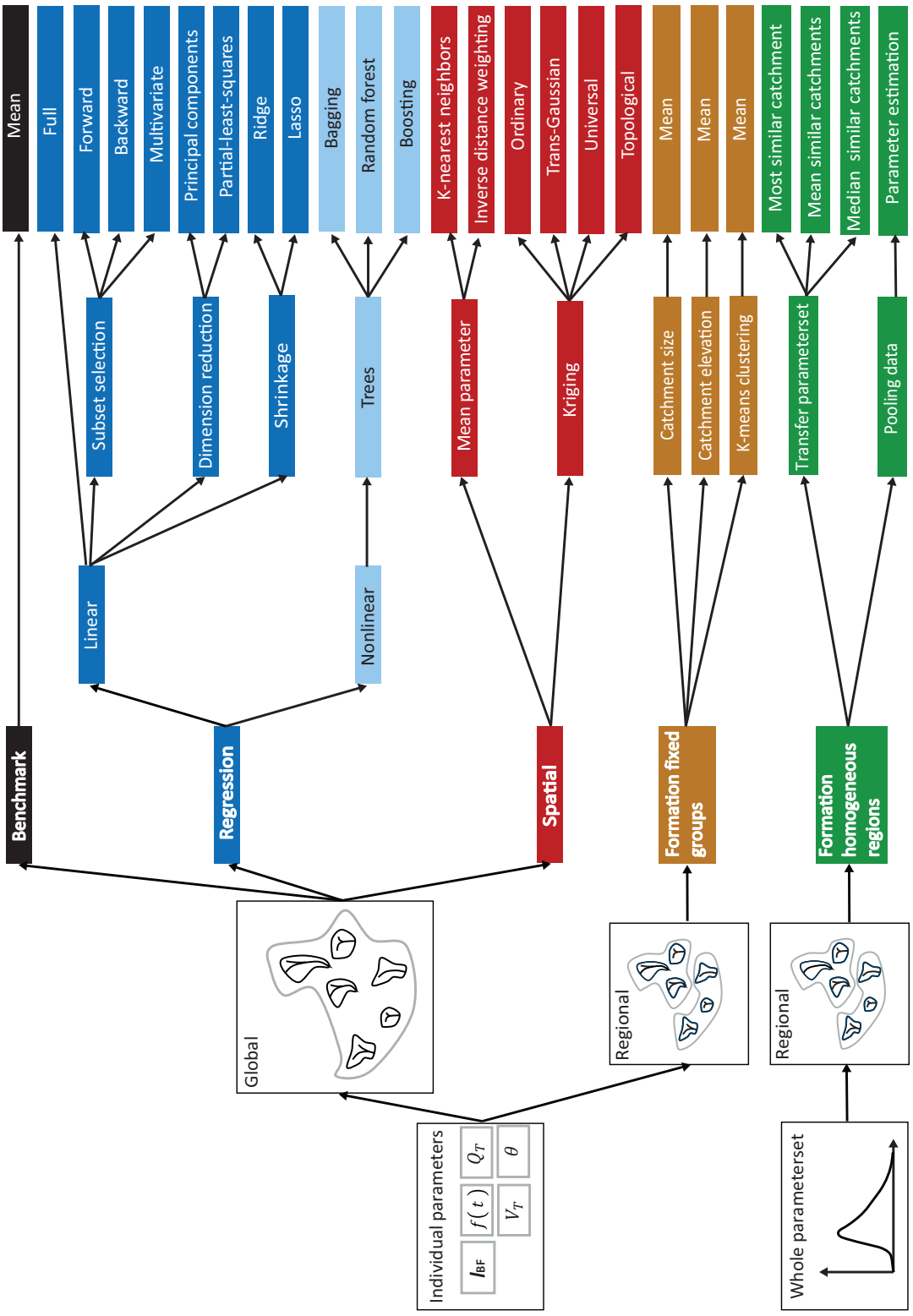



Figure 3.3: Methods tested to regionalize the ten parameters of a catchment-specific SDH on a global or regional scale. The methods were grouped into benchmark (black), regression (blue), and spatial (red) models, plus models that are based on the formation of regions using catchment characteristics (brown and green).

### 3.4.1 Nonlinear regression models

In the context of this thesis, the term nonlinear regression models is used for tree-based regression methods (*Strobl et al.*, 2009). Tree-based methods split the space of catchment characteristics used as explanatory variables into a number of regions containing catchments with similar SDH parameters. The set of splitting rules used to divide the space of explanatory variables can be summarized as a tree (*Hastie et al.*, 2008). Tree models are usually applied as an ensemble to avoid overfitting and improve prediction accuracy. Such ensemble methods comprise bagging, random forest, and boosting (*James et al.*, 2013). Bagging uses bootstrap techniques to draw random samples from the catchment set, builds regression trees for each of these samples, uses each of the trees to make predictions for the SDH parameter under consideration, and averages these predictions. Random forest is similar to bagging but considers only a subset of catchment characteristics to be used in each split of the tree model to decorrelate the individual models (*Breiman*, 2001). Boosted regression trees (*Freund and Schapire*, 1996; *Friedman*, 2001, 2002) work similarly but fit the regression trees iteratively focusing on catchments modeled poorly by the existing collection of trees (*Hofner et al.*, 2014).

### 3.4.2 Model validation

The predictive performance of the regionalization methods tested was compared to a benchmark model represented by the arithmetic mean (*Parajka et al.*, 2005; *Razavi and Coulibaly*, 2013) over the SDH parameters of the 163 study catchments (Section 3.1). The validation of the methods was done using 10-fold cross validation by dividing the dataset into 10 parts of equal size, estimating the SDH parameters using 9 out of these folds, and predicting the SDH parameters and constructing the SDHs for the catchments in the remaining fold (*James et al.*, 2013). The validation was on the one hand performed for the individual SDH parameters and on the other hand for the constructed SDHs in terms of their relative error. In the first case, each of the regionalized SDH parameters was compared to the corresponding SDH parameter estimated using runoff observations. In the second case, the hydrograph resulting from the regionalized SDH parameters was compared to the SDH estimated based on runoff observations. The relative error of the regionalized SDH compared to the estimated SDH was computed for four hydrograph characteristics (see Figure 3.4) including peak discharge ($Q_\mathrm{p}$), hydrograph volume ($V$), time to peak ($t_\mathrm{p}$), and half-recession time (the time from peak to where the recession reaches half of the peak discharge) ($t_{\mathrm{p}05}$).

## 3.5 SDH uncertainty

The uncertainty of constructed and regionalized SDHs was quantified in an uncertainty assessment framework. This framework is shortly outlined here and a detailed description is given in Paper IV. The framework consists of three levels of complexity (Figure 3.5). On a first level (A), the

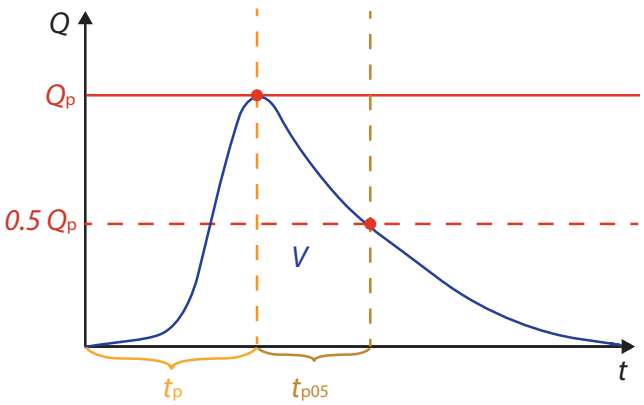



Figure 3.4: Illustration of the hydrograph characteristics used for the validation of the regionalization methods: peak discharge ($Q_p$), hydrograph volume ($V$), time to peak ($t_p$), and half-recession time ($t_{p05}$).

effect of different uncertainty sources comprising record length, model, and sampling uncertainty was assessed. This allowed for the identification of important uncertainty sources. On a second level (B), the total uncertainty of constructed SDHs and regionalized SDHs was assessed without considering their dependence. On a third level (C), the uncertainty of the constructed SDHs was propagated through regionalization to assess the coupled uncertainty of SDH construction and regionalization. The uncertainty assessment was based on bootstrap simulations which involve the random sampling of data points, with replacement, from the original sample (*Efron and Tibshirani*, 1993). The bootstrapping was done on various model configurations (model choices and parameters) to construct a set of SDHs. The resulting set of SDHs was compared to an SDH obtained as the best estimate under the standard configuration (see Section 3.3).

### 3.5.1 Level A: Uncertainty due to individual sources

Level A distinguished between three categories of uncertainty: 1) uncertainty due to limited record length, 2) model uncertainty resulting from the choice of one model over another feasible model, and 3) sampling uncertainty resulting from estimating the model parameters based on an available flood sample that only approximates the characteristics of the underlying population. The impact of the individual uncertainty sources on the estimated SDH was assessed by focusing on one uncertainty source at the time. This was done by constructing various SDHs using a slight modification of the standard model configuration (see Section 3.3) either varying one model choice or considering the uncertainty of one parameter at the time. The set of SDHs obtained by these simulations gave an idea of the variability introduced by each source of uncertainty. The uncertainty sources considered comprised record length, choice of the flood sampling strategy, parameter uncertainty in baseflow separation, choice of the depth function used to identify the median normalized hydrograph (RNH), choice of the PDF used to model the hydrograph shape, choice of the marginal distributions to model peak discharges and hydrograph volumes, uncertainty in the estimation of the parameters of the marginal distributions, choice of the copula

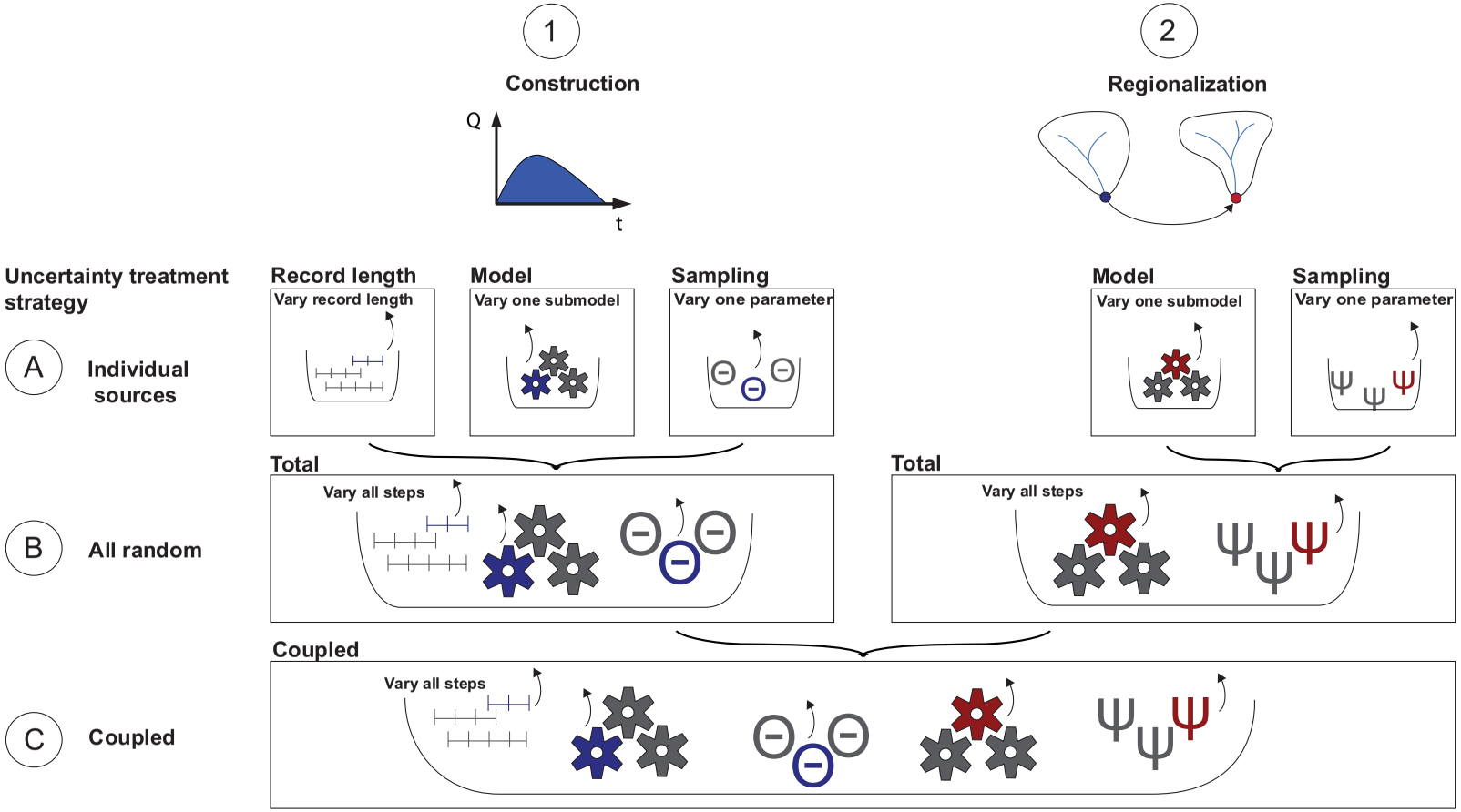



Figure 3.5: Uncertainty assessment framework with three levels of complexity: A) Uncertainty introduced by individual sources, specifically, record length, model choices, and sampling uncertainty; B) Total uncertainty of the constructed SDH and total uncertainty of the regionalized SDH resulting from the sources described in A; C) Coupled uncertainty of the SDH when construction uncertainty (steps 1A-1B) is propagated through regionalization (steps 2A-2B) onto the final regionalized SDH.

family to model the dependence between peak discharges and hydrograph volumes, uncertainty in the estimation of the parameter of the copula, choice of a design variable pair on the isoline, baseflow addition, and model and parameter uncertainty associated with regionalization.

### 3.5.2 Level B: Total uncertainty

The total uncertainty of SDH construction and SDH regionalization was assessed independently. The total uncertainty related to SDH construction was assessed by a bootstrap experiment. Contrary to Level A, neither model choices nor parameters were fixed. At each step of the construction procedure, one option was randomly sampled from the model and/or parameter space to jointly consider all uncertainty sources involved in SDH construction. The total uncertainty due to SDH regionalization was also assessed by bootstrap simulations. In each run, a regionalization model was sampled from the model space and fitted for each SDH parameter using data from a resampled catchment dataset. The fitted models were then used to predict the SDH parameters of the catchments in the original catchment set.

### 3.5.3 Level C: Coupled uncertainty

The coupled uncertainty jointly considered construction and regionalization uncertainty. Therefore, the uncertainty of the constructed SDHs was propagated through the regionalization process. Catchment-specific distributions of the ten SDH parameters were defined on Level B by constructing 1000 SDHs under various bootstrapped model configurations. This construction uncertainty was then propagated through regionalization by fitting a randomly sampled regionalization model to a resampled catchment set. While the ten SDH parameters were fixed for each catchment in the total regionalization uncertainty, the coupled uncertainty sampled the ten SDH parameters for each of the resampled catchments from their catchment-specific empirical distributions.

### 3.5.4 Uncertainty quantification

Each level of the uncertainty analysis provided a set of SDHs obtained by varying one or several model choices and/or considering sampling uncertainty. These sets were compared to the best estimate SDH obtained by the standard configuration (see Section 3.3) without considering uncertainty. The absolute relative error of each simulated SDH compared to the best estimate SDH was computed and summarized over all catchments as the median absolute relative error ($E_{\mathrm{MAR}}$). The $E_{\mathrm{MAR}}$ was computed for the four hydrograph characteristics $Q_{\mathrm{p}}$, $V$, $t_{\mathrm{p}}$, and $t_{\mathrm{p05}}$, which were already used for the validation of the regionalization approaches in Paper III (Figure 3.4).

## 3.6 Reactivity clusters

The comparison of approaches for the regionalization of catchment-specific SDHs in Paper III has shown that the hydrograph shape is difficult to regionalize since hydrograph shape variability is neglected. The regionalization of flood-type specific SDHs using the methods tested in Paper III was not found to be successful either. Therefore, the regionalization of event-type specific SDH sets was approached by a regional procedure. This regional procedure first identified representative sets of hydrograph shapes on a catchment scale. These sets were used in a second step for the delineation of regions of catchments with a similar flood reactivity. In a next step, a classification rule was defined that allowed for the attribution of an ungauged catchment to one of the reactivity regions. In a last step, the pooled data from the reactivity regions was used in a regional procedure to construct representative sets of synthetic design hydrographs for ungauged catchments. The first two steps, the identification of representative hydrograph sets, and the delineation of regions with a similar flood reactivity were addressed in Paper V and are summarized in this paragraph. The last two steps, the attribution of an ungauged catchment to one of the reactivity regions and the regional SDH construction procedure are addressed in Paper VI and outlined in Section 3.7.

The information included in hydrograph shapes was employed for the identification of representative hydrograph shapes on a catchment scale and for the delineation of regions with a similar flood behavior on a regional scale. Both representative hydrograph shapes and reactivity regions were identified using a clustering approach based on hydrograph shapes represented as functional data (FD) (Figure 3.6).

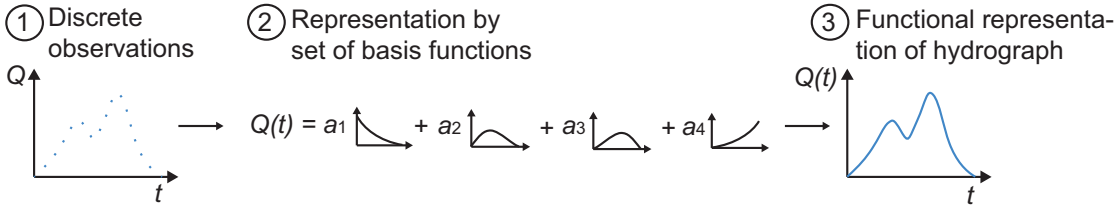



Figure 3.6: Getting from discrete measurements (1) to a functional representation of a hydrograph (3) by representing the data by a set of basis functions (2).

FD are continuously defined and do not depend on the choice of several hydrograph characteristics as do classical multivariate data. The clustering approach consisted of two steps: 1) identification of a set of representative hydrograph shapes on a catchment scale and 2) delineation of regions of catchments with a similar flood reactivity on a regional scale using the sets of representative hydrograph shapes obtained in the first Step (Figure 3.7). In the first Step, a clustering approach was used to identify the number of hydrograph shapes necessary to sufficiently describe the variability of hydrograph shapes within a study catchment. Hydrograph shapes were expressed as FD, i.e., each hydrograph was represented as a function of time (Figure 3.6). The FD representation was achieved by projecting the normalized hydrographs of a catchment on a set of

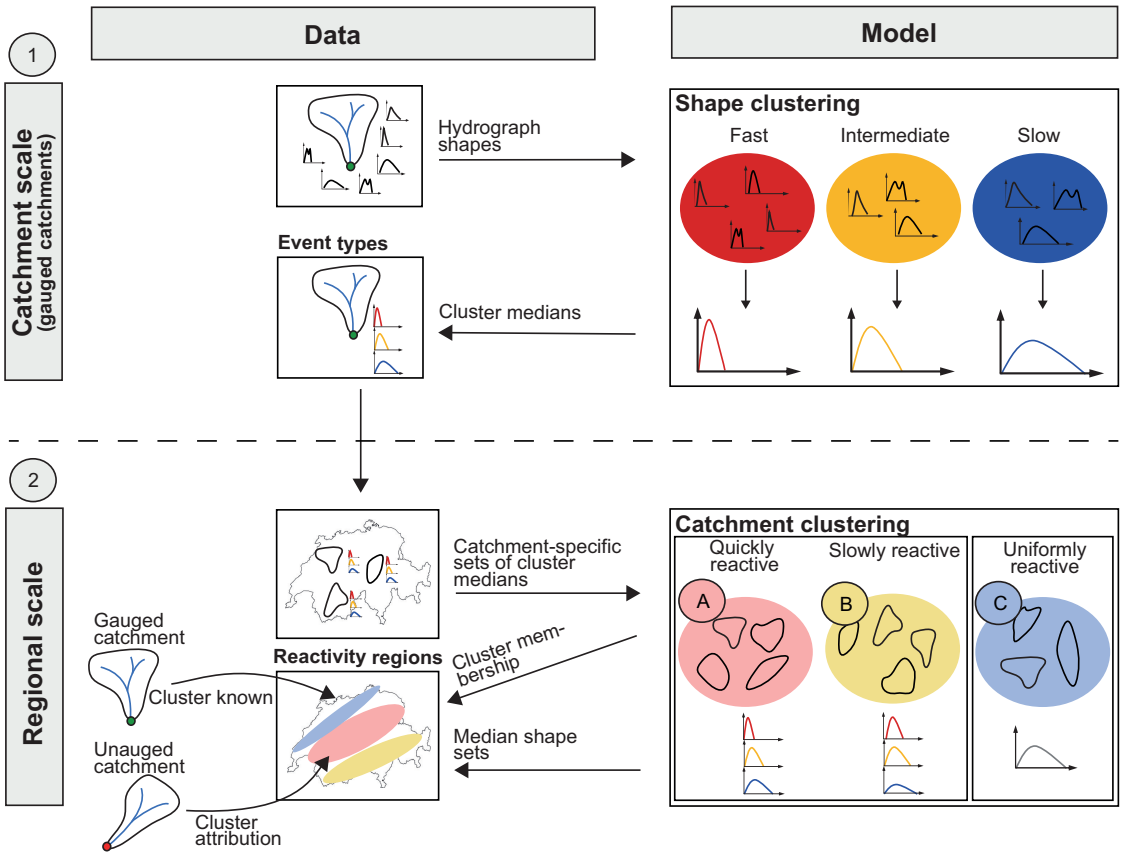



Figure 3.7: Illustration of the clustering framework. The data input and output for the models are indicated for the two different parts: 1) hydrograph shape clustering, 2) catchment clustering. Hydrograph shape clustering is done on the catchment scale and catchment clustering on the regional scale.

basis or B-spline basis functions and computing the coefficients for each of these B-spline bases (*Abraham et al.*, 2003). The sets of coefficients representing the hydrograph shapes were used as an input for the cluster analysis using a $k$-means algorithm (*Gordon*, 1999). Each hydrograph shape cluster was summarized by its median hydrograph to define a catchment-specific set of representative hydrograph shapes representing typical event types within a catchment. In the second step, regions of catchments with a similar flood behavior were delineated using the sets of representative hydrograph shapes identified in the first step. The catchments were clustered using the hierarchical Ward algorithm (*Ward*, 1963) on the sets of representative hydrograph shapes obtained in the first Step which were again summarized by a set of coefficients for a set of B-spline bases. A more detailed description of the two-step procedure can be found in Paper V.

## 3.7 Regional SDH sets

The regional approach for the construction of event-type specific SDH sets built up on the flood reactivity regions delineated in Paper V and consisted of three main steps (Figure 3.8):

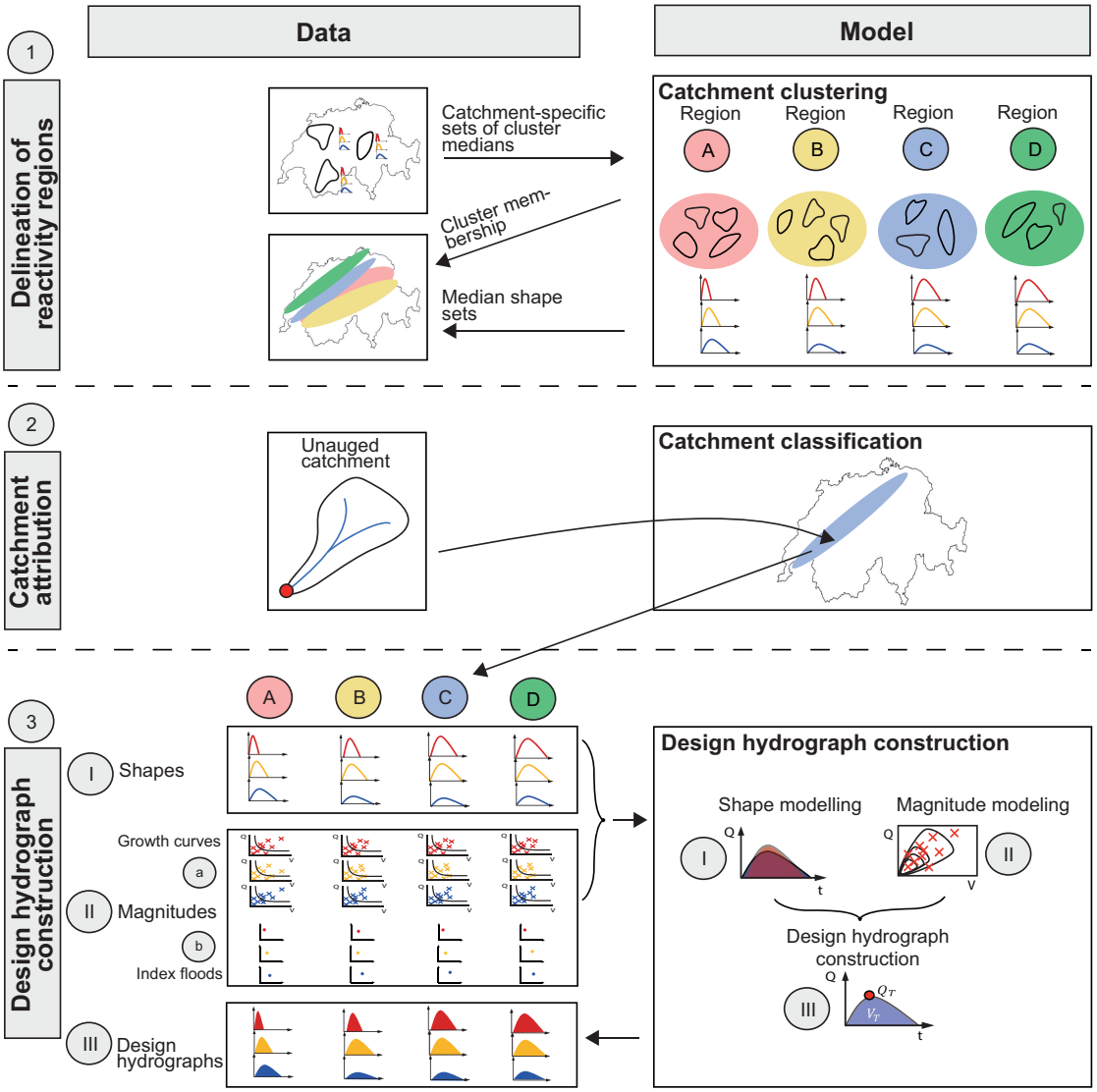



Figure 3.8: Regional event-type specific SDH construction framework: 1) Delineation of reactivity regions, 2) Catchment attribution, 3) Design hydrograph construction. Design hydrographs (III) characterize a flood both in terms of hydrograph shape (I) and hydrograph magnitude (II). The magnitude is estimated using event-type specific index floods (b) and regional growth curves (a).

1) delineation of reactivity regions, 2) attribution of an ungauged catchment to one of these regions using a classification rule, and 3) design hydrograph construction. These steps are shortly outlined below and explained in more detail in Paper VI.

The functional clustering approach proposed in Paper V was applied to establish four instead of three regions with a similar flood reactivity. A random forest model was then fitted to the region memberships of the catchments in the dataset and their catchment characteristics. The use of this model allowed the prediction of the best class membership of an ungauged catchment to one of the four reactivity regions. In addition, it allowed the prediction of probabilistic class memberships

for each of the reactivity regions. Both the best class membership and the probabilistic class memberships were used to construct SDHs for ungauged catchments by applying a bivariate index flood approach (*Requena et al.*, 2016). We first focus on the approach using the best class membership. The event-type specific and dimensionless data of the catchments belonging to the region that the ungauged catchment was assigned to with the highest probability of membership was pooled, i.e., a pool of fast events, a pool of intermediate events, and a pool of slow events was formed. Each pool consisted of event-type specific peak discharges, hydrograph volumes, and a representative hydrograph shape. The pooled dimensionless data were used in a bivariate regional frequency analysis to estimate event-type specific bivariate regional growth curves. In parallel, generalized linear models, which were fitted using observed event-type specific index floods, were used to predict the event-type specific index floods of the catchment under consideration. The event-type specific growth curves could then be used together with the index floods to estimate event-type specific design variable quantiles. These were used together with the corresponding representative hydrograph shapes fitted by a PDF to construct event-type specific sets of SDHs in ungauged catchments. Instead of basing regional SDH set construction on the pool of data of the region that was predicted with the highest probability, the probabilistic class memberships to each of the reactivity regions were used to compute weighted SDHs using data from all the regional data pools.

CHAPTER 4

# RESULTS

This chapter provides a summary of the results obtained by the analyses conducted within the scope of this thesis. Detailed results are presented in the six papers which can be found after the summary text.

## 4.1 Bivariate return periods

Paper I showed that different return period definitions, whose choice depends on the problem at hand, resulted in different estimated bivariate design variable quantiles, as illustrated on the study catchment Birse at Moutier-la-Charrue in the Swiss Jura (Figure 4.1). The conditional, joint, and Kendall's approaches resulted in isolines representing bivariate design variable quantiles of the same return period. The most probable design realization can be selected if a single design variable pair is of interest. The joint OR approach resulted in higher design variable quantiles compared to the univariate quantile. In contrast, the choice of the conditional approaches, the joint AND approach, the Kendall's approach, and the survival Kendall's approach resulted in lower design variable quantiles than the univariate case. These results demonstrated that the choice of the return period definition has a significant influence on the outcome of the design variable quantiles. It is therefore essential to well define the problem at hand to make a suitable choice of a return period definition. The case study highlighted that a univariate analysis can not provide a complete assessment of the probability of occurrence of a flood event if two or more dependent variables are of importance in the design process.

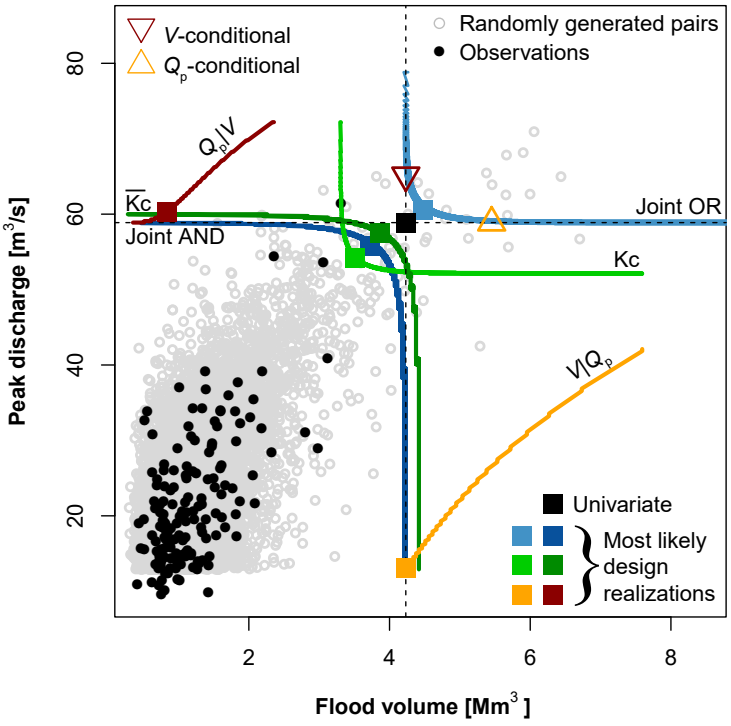



Figure 4.1: Design variable quantiles for different return period definitions. The black dots represent the observed flood events for the Birse catchment at Moutier-la-Charrue and the gray dots are 10 000 randomly generated pairs using the bivariate distribution of the peak discharges and hydrograph volumes. The black square represents the univariate design variable pair. The triangles represent the design variable pairs resulting from the $Q_p$- and $V$-conditional approaches applied to the joint OR isoline. The isolines represent the return level curves for the two joint approaches AND and OR, and the approaches using the Kendall's ($K_c$) and survival Kendall's ($\bar{K}_c$) distribution functions. The squares on the isolines represent the most-likely design realizations on these isolines.

## 4.2 SDH construction

The flood-type specific construction of SDHs as introduced in Paper II (see Section 3.3) has shown that different flood types are characterized by different SDH parameters. The parameters for the baseflow index, dependence between peak discharge and hydrograph volume, and the location and scale parameters of the lognormal PDF were dependent on the flood type while the parameters of the marginal distributions of the peak discharges and hydrograph volumes did not show any dependence on the flood type. The differences in some of the parameters resulted in different SDHs for each flood type as illustrated by three example catchments (Figure 4.2): Langete at Huttwil (a), Mentue at Yvonand (b), and Birs at Münchenstein (c). The specific peak discharges (peak discharge per unit area) only slightly differed when looking at catchments of all sizes. On the contrary, specific peak discharges of different flood types clearly differed within groups of catchments of similar size. In small catchments (20 - 75 km$^2$), short events such as FFs and SRFs

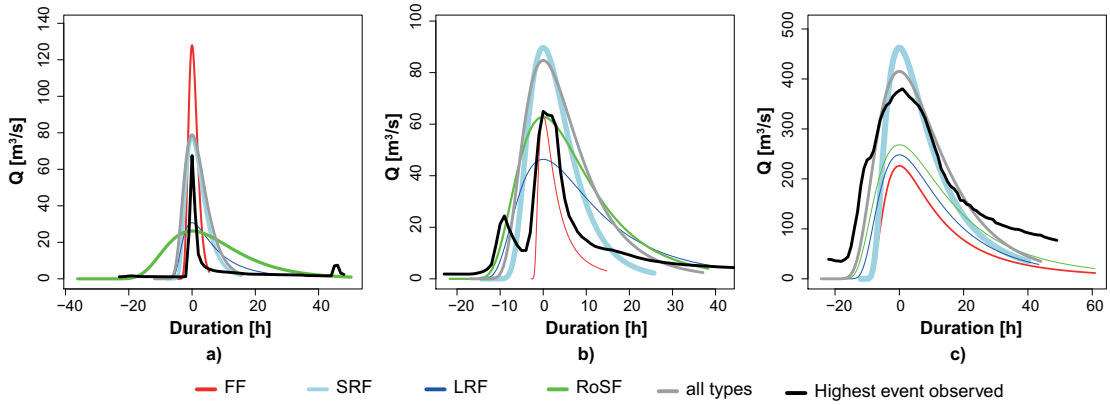



Figure 4.2: Flood-type specific SDHs for three catchments of the same mean elevation zone (650-800 m.a.s.l.) but different sizes (a) Langete at Huttwil: 60 km$^2$; b) Mentue at Yvonand: 105 km$^2$; c) Birs at Münchenstein: 911 km$^2$. The duration is centered around the time of occurrence of the peak which was set to zero and therefore the time is negative before and positive after the peak. The line width of the SDH represents the frequency of occurrence of a certain type in the respective catchment. The highest observed event in the catchment is shown in black.

showed generally higher specific peak discharges than longer events such as LRFs and RoSFs. The contrary was observed in medium size (76 - 300 km$^2$) and large catchments (301 - 1700 km$^2$), where LRFs and RoSFs were generally characterized by higher specific peak discharges than FFs and SRFs. Specific flood volumes were found to be higher for LRFs than RoSFs and clearly higher than for SRFs and FFs independent of the catchment size.

## 4.3 SDH regionalization

The comparison of 24 regionalization methods performed in Paper III included linear and nonlinear regression techniques, spatial methods, and methods based on the delineation of homogeneous regions (see Section 3.4). It showed that a better model than the benchmark model can be found for predicting catchment-specific SDHs in terms of the absolute relative error ($E_{\mathrm{AR}}$) (Figure 4.3). However, not all SDH parameters could be predicted equally well. The parameters $I_{\mathrm{BF}}$, $\theta$, and the parameters of the marginal distributions of peak discharges and hydrograph volumes could be regionalized but no suitable regionalization model could be found for those SDH parameters describing the hydrograph shape (PDF location and scale). In general, nonlinear regression techniques (especially boosted regression trees) performed slightly better than linear regression methods.

Different catchment characteristics were important for the prediction of the individual SDH parameters. Catchment characteristics related to geology and hydrogeology were important for the prediction of the $I_{\mathrm{BF}}$. Geology was also important for the prediction of the SDH parameters related to the shape of the hydrograph. On the contrary, catchment area was important for the prediction of the SDH parameters related to the event magnitude. Exposition was meaningful

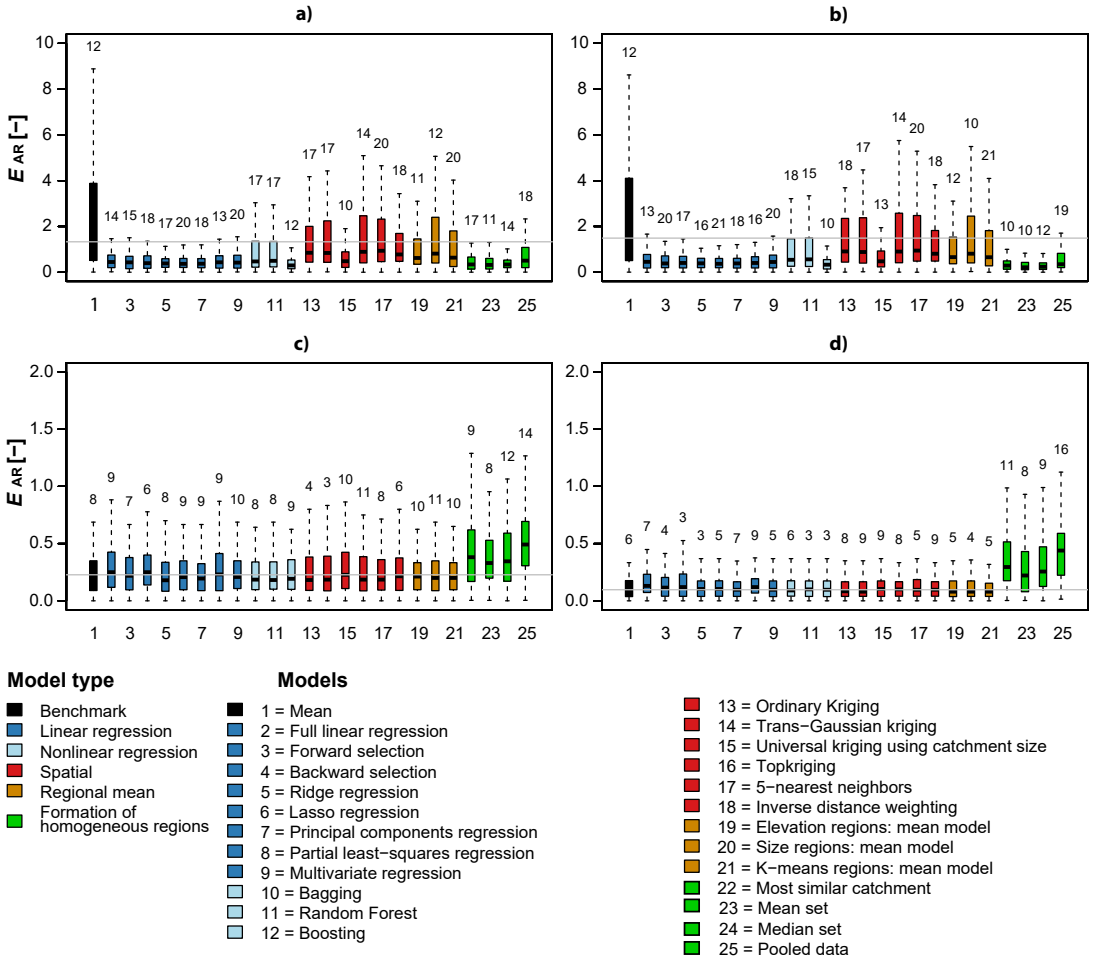



Figure 4.3: Predictive performance of the different regionalization methods tested for the different hydrograph characteristics a) peak discharge ($Q_p$), b) hydrograph volume ($V$), c) time to peak ($t_p$), and d) half-recession time ($t_{p05}$) provided as boxplots of the absolute relative error ($E_{AR}$) for the 163 study catchments. The number of outliers (defined as those observations lying outside the quartile ± 1.5 times the interquartile range) is indicated by the numbers plotted above the boxplots.

for the prediction of the shape parameters of the marginal distributions of peak discharges and hydrograph volumes and for the prediction of their dependence parameter $\theta$.

In contrast to catchment-specific SDHs, no suitable regionalization method could be found for the regionalization of flood-type specific SDHs as proposed in Paper II. This might have been related to the fact that the flood types employed in Paper II were not distinct enough.

## 4.4 SDH uncertainty

Table 4.1 summarizes the results of the uncertainty assessment conducted in Paper IV for the three levels of complexity A to C, where level A focused on individual uncertainty sources, level B on total construction and regionalization uncertainty, and level C on their coupled uncertainty.

Table 4.1: Summary of uncertainties across the three levels of complexity for the four hydrograph characteristics $Q_p$, $V$, $t_p$, and $t_{p05}$. The uncertainties are provided in the form of the 1st, 2nd, and 3rd quartile of the $E_{MAR}$s over all catchments. The numbers were rounded to two decimals.

| | $Q_p$ | | | $V$ | | | $t_p$ | | | $t_{p05}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **A: Uncertainty sources** | **1st** | **2nd** | **3rd** | **1st** | **2nd** | **3rd** | **1st** | **2nd** | **3rd** | **1st** | **2nd** | **3rd** |
| Sample size | 0 | 0.03 | 0.06 | 0.02 | 0.09 | 0.19 | 0 | 0.19 | 0.5 | 0 | 0.17 | 0.45 |
| Sampling strategy | 0.05 | 0.08 | 0.17 | 0.08 | 0.12 | 0.2 | 0.11 | 0.2 | 0.27 | 0.11 | 0.17 | 0.24 |
| RNH definition | 0 | 0 | 0 | 0.01 | 0.03 | 0.04 | 0.07 | 0.16 | 0.23 | 0.01 | 0.02 | 0.03 |
| PDF choice | 0 | 0 | 0 | 0.03 | 0.04 | 0.06 | 0.24 | 0.27 | 0.32 | 0.04 | 0.06 | 0.09 |
| Copula choice | 0 | 0.01 | 0.01 | 0.01 | 0.02 | 0.03 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.02 |
| Choice on isoline | 0.02 | 0.03 | 0.04 | 0.04 | 0.05 | 0.07 | 0.07 | 0.09 | 0.11 | 0.07 | 0.09 | 0.11 |
| Margin $Q_p$ | 0.12 | 0.2 | 0.34 | 0 | 0 | 0 | 0.09 | 0.14 | 0.2 | 0.09 | 0.14 | 0.2 |
| Margin $V$ | 0 | 0 | 0 | 0.08 | 0.12 | 0.16 | 0.08 | 0.12 | 0.16 | 0.08 | 0.12 | 0.16 |
| Baseflow separation | 0.03 | 0.04 | 0.06 | 0.07 | 0.11 | 0.16 | 0.09 | 0.13 | 0.21 | 0.1 | 0.15 | 0.2 |
| Baseflow addition | 0.06 | 0.08 | 0.1 | 0.06 | 0.08 | 0.1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Copula parameter | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Margin $Q_p$ parameter | 0.08 | 0.1 | 0.14 | 0 | 0 | 0 | 0.07 | 0.1 | 0.14 | 0.07 | 0.1 | 0.14 |
| Margin $V$ parameter | 0 | 0 | 0 | 0.13 | 0.15 | 0.19 | 0.13 | 0.15 | 0.19 | 0.13 | 0.15 | 0.19 |
| **B: Total construction** | 0.21 | 0.26 | 0.35 | 0.28 | 0.34 | 0.4 | 0.41 | 0.45 | 0.52 | 0.37 | 0.44 | 0.5 |
| Regionalization model | 0.23 | 0.36 | 0.63 | 0.21 | 0.34 | 0.76 | 0.12 | 0.21 | 0.34 | 0.19 | 0.28 | 0.41 |
| Regionalization sampling | 0.11 | 0.15 | 0.2 | 0.12 | 0.16 | 0.22 | 0.18 | 0.22 | 0.3 | 0.18 | 0.22 | 0.3 |
| **B: Total regionalization** | 0.26 | 0.4 | 0.67 | 0.29 | 0.43 | 0.68 | 0.14 | 0.25 | 0.41 | 0.2 | 0.3 | 0.44 |
| **C: Coupled** | 0.47 | 0.59 | 0.94 | 0.45 | 0.55 | 0.66 | 0.43 | 0.48 | 0.54 | 0.44 | 0.48 | 0.53 |

**Level A: Uncertainty due to individual sources:** The median uncertainty due to the individual uncertainty sources (see Section 3.5) over all catchments ranged between 0% and 30% for a design event of a return period of $T$ = 100 years (Table 4.1). Besides the sample size, the sampling strategy was one of the uncertainty sources most strongly affecting the median absolute relative error ($E_{\mathrm{MAR}}$) of all four hydrograph characteristics (see Figure 3.4).

**Level B: Total uncertainty:** The total construction uncertainty for the four hydrograph characteristics was larger than each of the individual uncertainty sources. The largest uncertainty over all catchments was found for the hydrograph characteristics related to the shape of the hydrograph. Also the total regionalization uncertainty was larger than the sampling and model uncertainties considered individually. The largest uncertainty over all catchments was found for the hydrograph characteristics related to the magnitude of the hydrograph.

**Level C: Coupled uncertainty:** The uncertainty assessed via the coupled uncertainty strategy lay around 50% when looking at the median over all catchments. It was slightly lower for the characteristics related to the shape of the hydrograph than for those related to the magnitude.

## 4.5 Reactivity clusters

The clustering of hydrograph shapes within a catchment, as performed in Paper V, represented hydrograph shapes as FD and resulted in three representative hydrograph shapes or event types: fast, intermediate, and slow events. The fast events were characterized by rather steep rising and falling limbs while the slow events were characterized by elongated rising and falling limbs. The intermediate events lay somewhere in between by showing quite steep rising but rather flat falling limbs (Figure 4.4). The clusters established via the hydrograph shapes were shown to have a meaning in terms of event magnitude. Events in the fast event cluster were generally characterized by high peak discharges but low hydrograph volumes while events in the slow event cluster showed high volumes but low peak discharges. The magnitudes of intermediate events lay in between those of the fast and slow events. The three median hydrographs of the fast, intermediate, and slow event clusters built the catchment-specific set of representative hydrograph shapes. This set consisted of three distinct hydrograph shapes in most catchments. However, there were catchments where the fast, intermediate, and slow events were not distinct from each other and the runoff reaction was rather uniform. Such catchments built the region of uniformly reactive catchments. The remaining catchments showed different types of runoff reactions. They were clustered into quickly reactive and slowly reactive catchments according to their sets of representative hydrograph shapes (i.e. fast, intermediate, and slow hydrograph) since a fast hydrograph in one catchment did not necessarily represent a fast hydrograph in another catchment. At a regional level, catchment-specific sets of representative hydrographs were distinct between catchments of a generally quick and a generally slow flood runoff reaction.

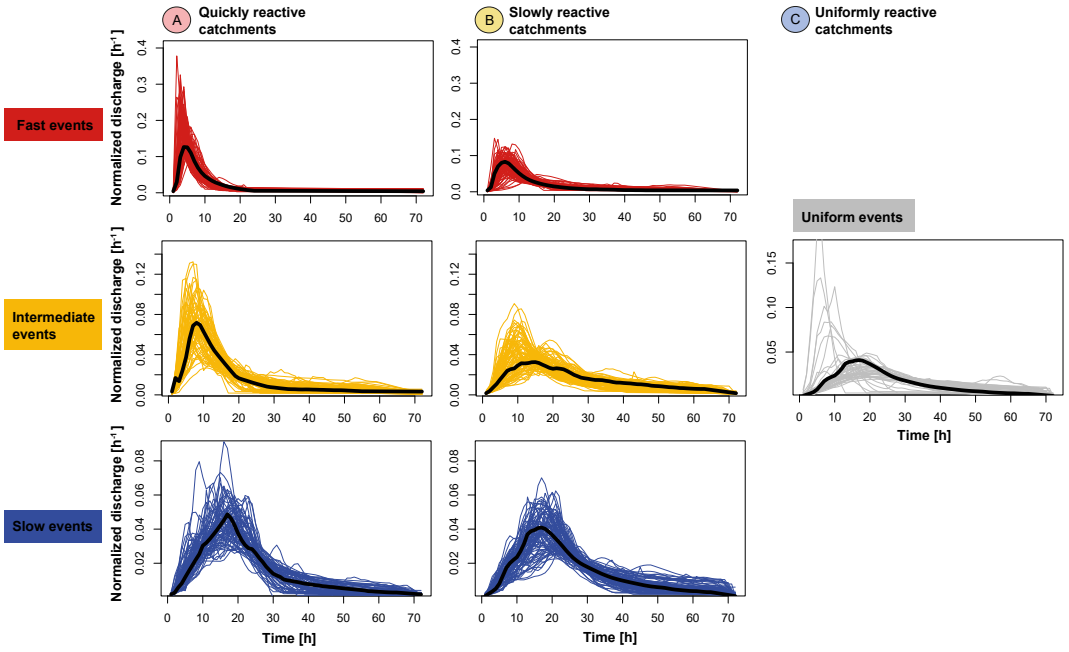



Figure 4.4: Representative hydrograph shapes of all catchments in a reactivity region (quickly reactive (A), slowly reactive (B), and uniformly reactive (C)). The median hydrograph shapes per region and event class (fast (red), intermediate (orange), and slow (blue)) are indicated in black. The scales of the $y$-axes are the same per event type but not across event types.

The quickly reactive catchments were mainly located in the Swiss Plateau, slowly reactive catchments in the Jura mountains, the Alpine region, and the Swiss Plateau, and uniformly reactive catchments mainly in the Jura mountains (Figure 4.5). The reactivity regions not only

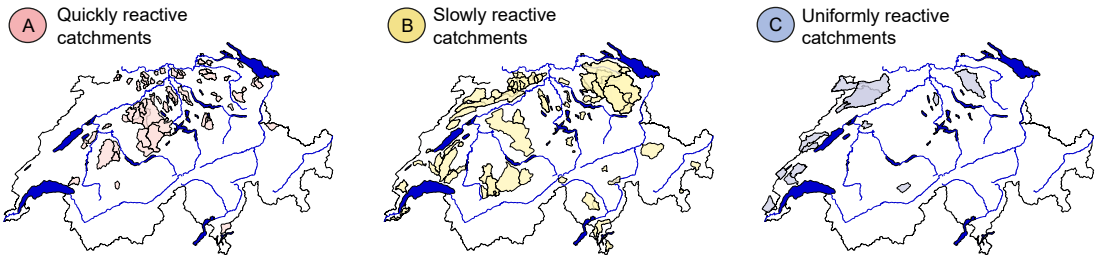



Figure 4.5: Catchments belonging to the quickly reactive (A), slowly reactive (B), and uniformly reactive (C) regions.

differed in terms of their representative hydrograph shapes but also in terms of their hydrograph magnitudes. Events occurring in the quickly reactive catchments were characterized by rather high peak discharges compared to hydrograph volumes while events occurring in the slowly reactive catchments showed rather high volumes compared to peak discharges. The events occurring in the three regions and belonging to the three event types not only showed different runoff characteristics (magnitudes and shapes) but differed also in their pre-event conditions in

the form of their triggering precipitation events (total amount and intensity).

## 4.6 Regional SDH sets

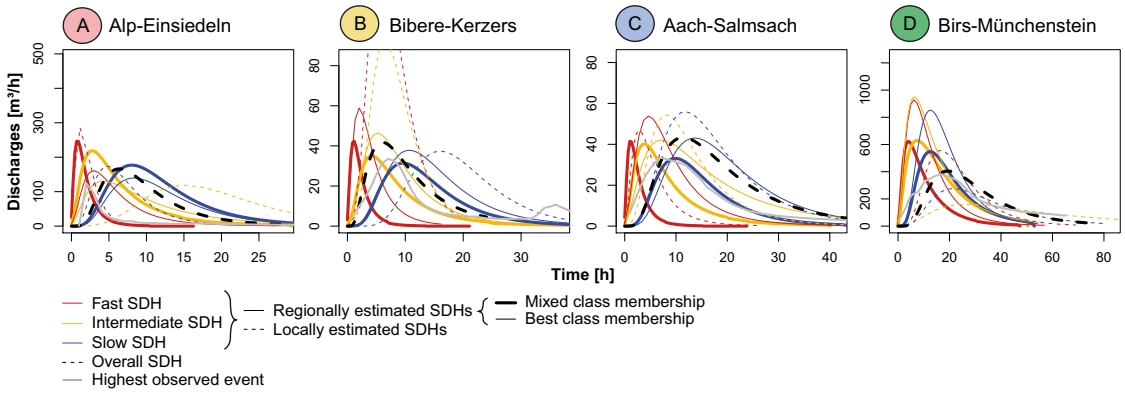



Figure 4.6: Event-type specific SDHs for four example catchments, one out of each reactivity region: Quickly reactive: A. Alp-Einsiedeln, intermediately reactive: B. Bibere-Kerzers, slowly reactive: C. Aach-Salmsach, uniformly reactive: D. Birs-Münchenstein. The regionally estimated event-type specific SDHs (bold lines) are plotted together with the locally estimated event-type specific SDHs (dashed lines). The catchment specific SDH not distinguishing between the event types is plotted in black. The highest observed event is given in gray. Regional estimates were derived based on the best region membership (thin lines) and using probabilistic class memberships (thick lines).

The use of the regional SDH set construction approach in Paper VI allowed for the estimation of event-type specific SDH sets in ungauged catchments (for four examples see Figure 4.6). Event-type specific SDH sets were computed using both the best class membership and probabilistic class memberships. Regionally estimated SDHs were comparable to their locally estimated counterparts and the highest observed event and the event-type specific estimates mostly surrounded the catchment-specific local SDH estimate. However, there were catchments for which regional predictions were rather difficult. The deviation of the regional SDH set computed using the best class membership from the regional SDH set computed using probabilistic memberships depended on the weight of the best class membership region compared to the weights of the other regions.

CHAPTER 5

# DISCUSSION

This thesis demonstrates how SDHs can be constructed in gauged catchments. It shows how SDHs can be regionalized to ungauged catchments and how their uncertainty can be assessed. Contrary to existing approaches, the construction of SDHs is based on observed runoff observations only and neither requires rainfall data nor the modeling of the interaction between rainfall and antecedent conditions, which makes the approach easily applicable by practitioners. The SDH construction procedure is not limited to the construction of catchment-specific SDHs but can be extended to an ensemble-based approach by constructing sets of SDHs for different flood or event types.

## 5.1 Bivariate return periods

A univariate flood frequency analysis of annual flood peaks may not be sufficient in situations where storage has a significant effect on flood attenuation. In such situations, flood duration and/or volume should be considered along with the peak flood discharge (*Chowdhary et al.*, 2011). This thesis highlighted that peak discharges and hydrograph volumes are dependent on each other and should jointly be considered in design flood estimation (Papers I-VI). Even though their joint consideration is beneficial from a theoretical point of view, several practical problems are associated to it. Among them, the definition of a return period in a bivariate setting. Paper I pointed out that the return period is no more uniquely defined in a multivariate setting when two design variables such as peak discharge and hydrograph volume are of interest. The case study presented in Paper I illustrated that the bivariate design quantiles depend on the choice of the return period definition. This highlighted the need of an accurate problem definition by the engineer or practitioner in order to choose a suitable return period definition. Even though this

has been stated in literature (*Gräler et al.*, 2013; *Serinaldi*, 2015), there is still a need to better outline which return period definition should be used in which design context and for what kind of design problems. For simplicity, this thesis focused on a joint OR return period assuming that both peak discharges and hydrograph volumes are equally important for the design problem at hand.

The case study in Paper I further showed that many design variable pairs with the same return period can be found in a multivariate setting. In practice however, it is often preferable to work with one design realization out of a set of pairs with the same return period. This implies the need to choose one design variable pair out of a set of pairs, which introduces some subjectivity. Such subjectivity could be avoided by moving from single-event based analyses towards ensemble-based analyses where a set of design variable pairs with the same return period is considered (*Gräler et al.*, 2013; *Salvadori et al.*, 2011). Alternatively, the frequency analysis could be moved to a later stage in a design analysis by not focusing on the flood events directly but on the final quantity of interest, e.g. water levels already accounting for reservoir routing (*Requena et al.*, 2013), flood damage or loss (*Apel et al.*, 2004; *de Moel et al.*, 2015; *Meyer et al.*, 2009), or the risk of failure (*Serinaldi*, 2015). The risk of failure has a unique definition independent of the nature of data and allows the consideration of both independent and dependent variables in stationary but also non stationary settings. A multivariate failure approach to assess hydrological risk in a general and consistent mathematical way seems valuable and has been outlined by *Salvadori et al.* (2016).

Bivariate frequency analysis is based on the same assumptions as univariate frequency analysis: stationarity, independence, and homogeneity (*Chebana and Ouarda*, 2008). Stationarity implies that the runoff regime has not been significantly altered by climate or land-use changes. The fulfillment of this assumption has been assured by carefully choosing the catchments included in the dataset used in this thesis. Catchments were chosen only if their runoff had not been altered by regulated lakes upstream or inland canals nor by urbanized areas. Climate induced changes are difficult to detect due to the length of the time series and the inter-annual variability in the data (*Bormann et al.*, 2011). However, numerous studies have identified significant trends in annual floods which questions the assumption of stationarity in certain case studies (*Šraj et al.*, 2016). The bivariate flood frequency estimation approach (*Gado and Nguyen*, 2016) and the definition of the return period (*Salas and Obeysekera*, 2014) as used in Papers I-IV and VI would need to be revised in a non stationarity context. In a bivariate framework, non-stationarity can emerge in the statistical attributes of the univariate variables, in the dependence structure of the variables, or both (*Bender et al.*, 2014).

The independence of events is generally easier to ensure if an annual maximum flood sampling approach is chosen instead of a peak-over-threshold approach. Still, this thesis used flood events sampled using a peak-over-threshold approach which allows the inclusion of all potentially relevant events (*Lang et al.*, 1999). Independence was ensured by prescribing a minimum time

window of 72 hours between successive events.

The assumption of homogeneity was not rejected by statistical tests, however, inhomogeneity could possibly be introduced by different flood types (*Szolgay et al.*, 2015). The flood-type specific frequency analysis introduced in Paper II alleviates this potential problem by considering each flood type as a separate sample. Non homogeneity could also be an issue when pooling events of different catchments (*Hosking and Wallis*, 1997) as done in some regionalization approaches (Papers II and VI).

## 5.2 SDH construction

The SDH construction approach (Paper II) is based on observed runoff data only and does not require any rainfall data. This distinguishes it from event-based and continuous approaches (*Grimaldi et al.*, 2012a) and makes it easy to apply by practitioners and engineers. There is neither a need to choose an event rainfall and specifying pre-event conditions as in an event-based approach nor the need of calibrating a hydrological model as in continuous approaches. However, the SDH construction approach hinges on the available data record and uncertainty can be quite large in the case of short records as shown in Paper IV. A combination of the proposed approach with a continuous scheme could therefore be beneficial. Long runoff time series could be simulated based on rainfall records, which are typically longer than runoff records, or using a stochastic model and be used for flood frequency analysis and SDH construction. However, the use of both a stochastic model and a rainfall-runoff model would introduce additional sources of uncertainty.

Paper II picked up the idea of an ensemble-based design approach introduced in Paper I, however, in a slightly different context than when considering a set of design variable pairs with the same bivariate return period. Single-event based SDH construction neglects process variability in a catchment. On the contrary, the construction of a set of flood-type specific SDHs allows for the consideration of different flood mechanisms. Flood events triggered by high-intensity precipitation events are represented by the flash flood SDH, events caused by short but less intense precipitation by the short-rain flood SDH, events caused by long lasting but little intense precipitation by the long-rain flood SDH, and events triggered by rainfall falling on snow by the rain-on-snow flood SDH. Instead of using flood type samples, one can use samples for the three representative event-types (fast, intermediate, and slow events) identified in Paper V to construct event-type specific hydrographs. Alternatively, seasonal SDHs could be constructed using seasonal flood samples as proposed by *Brunner et al.* (2017). An event-type specific SDH construction procedure allows for the representation of process variability within a catchment and is therefore a step forward from flood frequency statistics towards *flood frequency hydrology*, a framework introduced by *Merz and Blöschl* (2008b) fostering the expansion of information beyond the locally observed flood sample. A set of event-type specific hydrographs could be

used in hydraulic modeling to assess the effect of different hydrographs (with the same return period) on water levels or inundation areas which allows for a broad assessment of potential flood consequences. However, an event-type specific approach also has some drawbacks compared to a single-event approach. The size of the individual flood samples used for flood frequency analysis, one for each flood type, is significantly reduced. A small sample size leads to less reliable (*Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012) and more uncertain flood estimates as highlighted by the uncertainty assessment performed in Paper IV and several other studies (e.g. *Apel et al.*, 2004; *Beven and Hall*, 2014). The problem of a reduced sample size due to an event-type specific analysis could be relieved by pooling events of different flood types from similar catchments (*Grimaldi et al.*, 2016) as done in Paper VI. However, the identification of catchments with a similar behavior in terms of flood type occurrence is not straightforward. Different combinations of triggering precipitation mechanisms and pre-event catchment conditions might lead to similar flood events in terms of magnitude even though they differ in the flood type. The classification scheme used to define flood types based on several indices related to meteorology and catchment conditions was therefore not found to be suitable for the identification of regions with a similar flood behavior. It was found in Paper V that the identification of regions with a similar flood behavior should rather be based on the information stored in the hydrographs themselves which integrate pre-event conditions, precipitation characteristics, and runoff processes.

## 5.3 SDH regionalization

The comparison of various regionalization methods (Paper III) showed that a successful regionalization of catchment-specific SDHs is possible using nonlinear regression (especially boosted regression trees) and linear regression techniques. However, it also showed that the shape of the hydrograph is difficult to regionalize. This was related to the non-representativeness of a single median hydrograph shape for the different flood types observed in a catchment. A better representation of different potential shapes could be achieved via the flood-type specific design hydrographs proposed in Paper II. However, such flood-type specific SDHs were found to be difficult to regionalize, i.e., neither could a spatial pattern be detected, nor could a relationship between flood-type specific SDHs and catchment characteristics be established. There is no unique combination of rainfall and drainage basin condition which produces the design flood (*Merz and Blöschl*, 2009; *Pilgrim*, 1986), which might explain the problems encountered in establishing a relationship between SDH parameters and physiographical and meteorological catchment characteristics. *Merz and Blöschl* (2009) have pointed out the need for better indicators than static and climatological indicators such as runoff coefficients. They highlighted that such coefficients are not available in ungauged catchments and that new indicators at the regional scale have to be developed that are more representative of the causative flood processes. Therefore, Paper V focused on the hydrograph shapes observed in a catchment, which integrate pre-event conditions,

catchment characteristics, and meteorological input, which jointly influence the magnitude of an event, to identify regions of similar flood behavior. A catchment's flood behavior was summarized by a set of three representative event types comprising a fast event with both a steep rising and recession limb, an intermediate event with a steep rising but rather flat recession limb, and a slow event with both a flat rising and recession limb. Grouping the catchments according to their sets of representative event types allowed for the identification of three regions with catchments of a similar flood behavior. Such regions were shown to be pertinent for the regionalization of event-type specific SDHs to ungauged catchments in Paper VI. An ungauged catchment was attributed to one of the regions based on its characteristics using a classification rule. A regional index flood approach (*Requena et al.*, 2016) was then applied to event-type specific, pooled flood events of that region to construct event-type specific SDHs in an ungauged catchment. The use of these reactivity regions in combination with a bivariate index flood approach and an event-type specific SDH construction procedure similar to the one proposed in Paper II enabled the construction of an ensemble of SDHs. Similar to the set of flood-type specific SDHs constructed in gauged catchments, this event-type specific procedure for ungauged catchments allows for the representation of process variability in design flood regionalization.

## 5.4 SDH uncertainty

Several model choices are involved in the construction of SDHs in gauged catchments and their regionalization to ungauged catchments. Choosing one model among a set of feasible models and estimating the parameters of these models introduces uncertainty as shown in Paper IV. The uncertainty involved in SDH construction and regionalization should be communicated to the end-user (*Pappenberger and Beven*, 2006). One possibility to do so is a formal uncertainty analysis, as the one proposed in Paper IV, performed to derive uncertainty bands for constructed or regionalized SDHs. Another possibility would be to simplify such a procedure by focusing on the most important uncertainty sources identified in Paper IV, i.e., the record length, sampling strategy, and the choice of the marginal distributions for peak discharges and hydrograph volumes. An uncertainty source not addressed by such an approach is uncertainty introduced due to process variability observed within a catchment. This variability, could in gauged catchments, be depicted by a set of flood-type specific SDHs as outlined in Paper II or by a set of representative SDHs as outlined in Paper VI. The latter strategy can also be applied in ungauged catchments after having assigned a catchment to one of the flood reactivity regions (Paper VI).

The main source of uncertainty in SDH construction was found to be record length, which is propagated through all modeling steps including regionalization. It is not surprising that runoff data are the most valuable information since the observed runoff is an integrative indicator of the predominant hydrological processes in a catchment (*McGlynn et al.*, 2013). Uncertainty reduction should therefore start by expanding information on runoff. Three main categories of

information expansion have been suggested (*Merz and Blöschl*, 2008a): temporal, spatial, and causal information expansion. Temporal information expansion refers to the consideration of historical events that occurred before the period of systematic data collection. Such historical events are usually estimated based on documentary evidence. Most historical data on floods provide information on the water stage during peak discharge, which can be converted to peak discharge (*Brázdil et al.*, 2006; *Wetter*, 2017). However, information on hydrograph volume is hardly available. The use of a continuous rainfall-runoff model driven by rainfall series that are longer than observed runoff series our by rainfall series generated by a stochastic model would be a possibility for bivariate information expansion, jointly expanding peak discharge and volume series without neglecting their dependence.

Spatial information expansion refers to the use of data from similar/neighboring stations in a regional flood frequency analysis. An example of such an approach is the index flood approach (*Dalrymple*, 1960) which was extended to a bivariate setting (*Requena et al.*, 2016) and was used in Paper VI.

Causal information expansion refers to the expansion of information by including process-based information. This was done in Paper II by constructing flood-type specific SDHs and in Paper VI by constructing event-type specific SDHs in ungauged catchments. The use of the flood reactivity regions established in Paper V allowed for the pooling of floods belonging to the same event type (fast, intermediate, and slow events), which is also a type of causal information expansion. Information expansion using one way or the other helps to relieve the problem of choosing one model over the other and to reduce sampling uncertainty.

## 5.5 Reactivity clusters and regional SDH sets

The two-step clustering approach proposed in Paper V focused on hydrograph shapes and is a targeted approach for the identification of catchments similar in their flood behavior. It allowed first, for the identification of catchment-specific representative event types and second, for the identification of regions with a similar flood behavior based on these catchment-specific event types. The expression of the flood hydrographs as functional data enabled the use of all the information stored in the flood hydrograph, which is the most detailed signature of how a catchment behaves and the complex result of different processes (*Blöschl et al.*, 2013; *Hannah et al.*, 2000). Paper V showed that the flood hydrograph information can be used to identify regions of catchments with a similar flood behavior which is hardly possible when focusing only on static and climatological characteristics (*Merz and Blöschl*, 2009). The clusters were shown to have a hydro-meteorological meaning which makes them useful from a hydrological point of view (*Rosbjerg et al.*, 2013). The catchment classification is initially based on functions characterized by streamflow which makes it widely applicable (*Wagener et al.*, 2007). Paper V helps to overcome the difficulties of regionalizing hydrograph shapes as attempted in Paper III

where the hydrograph shapes within a catchment were only summarized and process variability was neglected. It provides a means of better representing process variability in design hydrograph construction by identifying event types characteristic of a catchment. The event type classification is, contrary to classical flood type classification schemes, such as the one applied in Paper II, independent of event and meteorological characteristics (*Merz and Blöschl*, 2003; *Sikorska et al.*, 2015).

Paper VI showed that ungauged catchments can be assigned to one of the reactivity regions via a classification rule using catchment characteristics. The construction of an event-type specific set of SDHs in an ungauged catchment can then be based on the pooled data of its predicted reactivity region. The event shapes can be modeled by the median fast, intermediate, and slow shape for that region. The event magnitude can be estimated using a bivariate index flood approach (*Requena et al.*, 2016) where the regional growth curve is estimated using the pooled, event type specific data from the region and the index flood is estimated based on catchment characteristics. It has been shown that flood quantiles obtained using a regional growth curve are more accurate than those obtained by at-site analysis (*Hosking and Wallis*, 1997; *Kysely*, 2008). The size of the pooled dataset allows a more reliable estimation of the shape parameter of the GEV or GPD distribution (*Evin et al.*, 2016) or the use of a more flexible distribution than these classical extreme value distributions with (only) three parameters (*Coles*, 2001). In paper VI, we used the flexible Wakeby distribution (*Hosking and Wallis*, 1997; *Houghton*, 1978; *Landwehr et al.*, 1978) with five parameters. Another potential candidate would be the four parameter Kappa distribution (*Hosking and Wallis*, 1997; *Mielke*, 1973).

## 5.6 Limitations and Perspectives

The approaches proposed in the context of this thesis can generally be applied to datasets other than the dataset used in this thesis. Such datasets could consist of catchments from another region with similar climatic conditions (temperate-humid), of catchments of different sizes, or of catchments with runoff records at another temporal resolution (e.g. daily instead of hourly if a daily resolution is able to resolve the main runoff features). However, the approaches should neither be applied in strongly anthropogenically influenced catchments since dams and reservoirs change hydrological variability and significantly affect hydrological extremes (*Di Baldassarre et al.*, 2017) nor in large areas where floods from different subcatchments superpose which would require the consideration of their spatial dependence (*Wang et al.*, 2014). In such cases, rainfall-runoff models might be applied (*Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012). However, the methods proposed could be applied on time series simulated with a rainfall-runoff model. The adaptation of the approaches to a new dataset requires the verification of several assumptions. First, it needs to be verified whether the assumptions of stationarity, homogeneity, and independence of events are fulfilled. Second, the SDH construction procedure

(Paper II) might need to be slightly adjusted regarding the typical event duration, the marginal distributions of peak discharges and hydrograph volumes, their dependence structure (copula), and the PDF used to model hydrograph shapes. Third, the choice of the return period definition needs to be made according to the design problem at hand (Paper I). Fourth, the suitability of the several regionalization procedures might need to be checked for the regionalization of catchment-specific SDHs to ungauged catchments (Paper III). Fifth, the uncertainty assessment framework might require an adjustment of model and parameter sets to sample from and/or different numbers of iterations (Paper IV). Sixth, the clustering approach might require adjustments in the number and kind of B-splines used to represent hydrograph shapes as functional data and in the number of event type and region clusters.

The construction of SDHs in gauged catchments and their estimation in ungauged catchments is currently limited to a return period of $T \leq 100$ since reliable estimates are only to be expected for events of a return period two or three times the available record length (*Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012). Still, more reliable estimates are possible if runoff information is expanded using historical events in both a maximum likelihood (*Stedinger and Cohn*, 1986) or a Bayesian framework (*Salinas et al.*, 2016). The work done in this thesis could be extended into several directions:

1. **Backward extension of modeling chain:** Runoff data simulated by a hydrological model could be used for SDH construction instead of observed runoff data. This would allow for an extension of the time series by using observed precipitation time series, which are usually longer than observed runoff time series, or by generating rainfall time series using rainfall generators (*Evin et al.*, 2016; *Vandenberghe et al.*, 2010).

2. **Forward extension of modeling chain:** Use of hydraulic models on catchment-specific or event-type specific SDHs for inundation modeling, risk mapping, or flood damage and loss modeling (*Merz and Thieken*, 2009; *Thieken et al.*, 2009). The use of event-type specific SDHs instead of catchment specific SDHs would allow for an assessment of the range of outcomes to be expected.

3. **Consideration of rating curve uncertainty:** Design flood estimation techniques often ignore the uncertainty in the underlying rating curve model even though rating curve uncertainty may play an important role when an extrapolation of the rating curve is necessary (*Steinbakk et al.*, 2016). The effect of rating curve uncertainty (*McMillan and Westerberg*, 2015) on bivariate design quantiles still needs to be assessed.

4. **Assessment of climate impact on bivariate design quantiles or SDHs:** The impact of climate change on bivariate design quantiles or SDHs could be assessed by slightly modifying the modeling chain classically used in climate impact studies consisting of an emission scenario, a global circulation model, downscaling, possibly including bias

correction, a hydrological model, and (univariate) flood frequency analysis (*Delgado et al.*, 2014). The last step in this chain could be replaced by a bivariate flood frequency analysis focusing not only on changes in peak discharges but also hydrograph volumes. Alternatively, it could be replaced by SDH construction which would allow for an assessment of changes in both the bivariate quantiles and hydrograph shapes (*Brunner et al.*, 2017).

5. **Assessment of past/future changes in flood/event types:** The development of different flood (Paper II) or event types (Paper V) over time could be assessed in a trend analysis similar to what has been done for peak discharges by *Hundecha et al.* (2017). Instead of going back in time, such an analysis could be done on time series simulated for future climate conditions (*Brunner et al.*, 2017).

6. **Assessment of effect of different flood types on flood damages/loss:** *Kreibich and Dimitrova* (2010) found that four flood types comprising floods caused by high groundwater levels, riverine floods, flash floods, and dyke breaches showed significantly different impact characteristics concerning the factors of water level, flood duration, flow velocity, contamination, and resulting loss. The effect of different flood types on flood loss could instead be assessed for the six flood types FFs, SRFs, LRFs, RoSFs, SMFs, and GMFs used in Paper II or for the three event types fast, intermediate, and slow defined in Paper V.

CHAPTER 6

# CONCLUSIONS

This thesis presents a simple and flexible modeling framework for the estimation of design floods both in gauged and ungauged catchments which enables the construction of catchment-specific and event-type specific SDHs. It allows for the consideration of different sources of uncertainty and for the representation of process variability.

The main conclusions can be summarized as:

- The use of a bivariate return period definition allows for the joint consideration of peak discharge and hydrograph volume in design flood estimation. However, the choice of this definition largely influences the magnitude of the design quantiles and a focus on a single-event approach requires the choice of one design event from an isoline of events with the same return period. Such bivariate design quantiles can be combined with a representative hydrograph shape modeled by a probability density function to form a synthetic design hydrograph (SDH).
- SDHs can be constructed using a specific subsample of flood events, e.g. subsamples describing different flood types or subsamples describing different event types. A flood-type or event-type specific SDH construction allows for the consideration of different process-types in design flood estimation and therefore for a better representation of process-variability than single-event approaches.
- Catchment-specific SDHs are best regionalized using boosted regression trees or linear regression methods. In contrast, the regionalization of event-type-specific SDHs requires the delineation of regions with a similar flood behavior and the assignment of an ungauged catchment to one of these regions. An SDH set can then be estimated using the pooled data from this region.

- The uncertainty of constructed and regionalized catchment-specific SDHs can be assessed via bootstrap simulations. The most important uncertainty sources in SDH construction are the record length and the choice of the flood sampling strategy. The coupled uncertainty of SDH construction and regionalization lies around 50%.

The approaches proposed within the framework of this thesis and the results obtained can be summarized by the following main contributions:

1. **From a univariate over a bivariate to a synthetic design hydrograph approach:** The thesis did not focus on univariate quantiles for peak discharges as do classical flood frequency analyses but developed tools to represent design floods as bivariate quantiles in terms of peak discharges and hydrograph volumes. These quantiles were combined with a representative hydrograph shape to complete the information provided by a design flood.

2. **From flood frequency statistics to *flood frequency hydrology:*** The methods and frameworks proposed in the six papers contribute to making a step from flood frequency statistics, which does not traditionally differentiate between flood event types and flood generation processes (*Merz et al.*, 2014), towards *flood frequency hydrology*, which considers that different flood producing processes imprint in a different way on the flood frequency curve (*Merz and Blöschl*, 2008b). Local flood data is combined with additional information such as spatial information on floods in neighboring catchments (Papers III, V, and VI), and causal information on the flood processes (Papers II, V, and VI). Spatial information is used in Paper III where some of the regionalization approaches tested use data from similar catchments to transfer SDH parameters from gauged to ungauged catchments. Causal information is used in Paper II where SDHs are constructed for four flood types representing different processes and in Paper V where catchment-specific sets of representative hydrograph shapes are used for the delineation of regions with similar flood behaviors.

3. **From a single-event to an ensemble design flood approach:** The classical design approach representing a design flood by a single event is extended to an ensemble approach where sets of design floods are constructed (flood-type specific or event-type specific SDHs) to better represent process variability within a catchment. The use of such design flood sets helps to illustrate the real threat to hydraulic structures (*Gräler et al.*, 2013) and to simulate impacts due to possible failures (*Schumann et al.*, 2010), allows for cost-benefit analyses (*Plate*, 2002), facilitates the shift from a safety oriented planning towards risk awareness (*Nijssen et al.*, 2009), and represents variability (*Klein et al.*, 2010) and uncertainties (*Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012).

4. **From gauged to ungauged catchments:** This thesis shows ways of regionalizing catchment-specific SDHs and ways of regionalizing event-type specific SDH sets to ungauged catch-

ments. The latter allow the representation of process variability even in ungauged catchments.

Altogether, this work encourages the use of sets of SDHs instead of a single SDH because SDH sets better represent process variability within a catchment.

# ACKNOWLEDGEMENTS

I would like to thank several persons who gave me advice and support during my time as a PhD student:

1. My two main supervisors Anne-Catherine FAVRE and Jan SEIBERT for their support throughout the thesis, many stimulating discussions, critical but constructive feedback on ideas and paper drafts, and their time dedicated to the project. Anne-Catherine for her helpful statistical explanations and questions, supporting me in improving my French, her motivating words and positive spirit, and her help in dealing with the French bureaucracy. Jan for his hydrological input, passing on his experiences as an editor and reviewer, and sharing the fascination for running and cross-country skiing.

2. My third PhD-committee member Reinhard FURRER for his pragmatic view, support and ideas regarding statistical problems, and his helpful explanations on both his black- and whiteboard.

3. Olivier VANNIER for introducing me to the topic of SDH construction at the very beginning of my thesis. My co-supervisors Anna SIKORSKA and Daniel VIVIROLI for helpful discussions, constructive inputs and feedback, reading and correcting paper drafts and reports, and for the good collaboration.

4. The Federal Office for the Environment (FOEN) for financing the project and the members of the advisory group, namely, Jean-Pierre JORDAN, Olivier OVERNEY, Martin BARBEN, Armin PETRASCHECK, Paul MEYLAN, Ion IORGULESCU, and Manfred STÄHLI for fruitful discussions and useful feedback.

5. Swissuniversities and Campus France for supporting the mobility between the Universities of Zurich and Grenoble via a grant.

6. Tracy EWEN for language editing the papers written in the context of this thesis.

7. My colleagues from the Hydrology and Climate group (H2K) in Zurich and the HMCI group in Grenoble for sharing experiences and exchanging thoughts.

8. My family for their support and open ears and Silvan for providing a different view, distracting me from problems, and being there when I need him.

# References


Abraham, C., P. A. Cornillon, E. Matzner-Lober, and N. Molinari (2003), Unsupervised Curve Clustering using B-Splines, *Scandinavian Journal of Statistics*, *30*, 1–15.

Apel, H., A. H. Thieken, B. Merz, and G. Blöschl (2004), Flood risk assessment and associated uncertainty, *Natural Hazards and Earth System Science*, *4*(2), 295–308, doi: 10.5194/nhess-4-295-2004.

Bender, J., T. Wahl, and J. Jensen (2014), Multivariate design in the presence of non-stationarity, *Journal of Hydrology*, *514*, 123–130, doi: 10.1016/j.jhydrol.2014.04.017.

Beven, K., and J. Hall (2014), *Applied uncertainty analysis for flood risk management*, 672 pp., Imperial College Press, London.

Beven, K., D. Leedal, and R. Alcock (2010), Uncertainty and Good Practice in Hydrological Prediction, *VATTEN*, *66*(3-4), 159–163.

Bezzola, G. R., and C. Hegg (2007), Ereignisanalyse Hochwasser 2005. Teil 1 - Prozesse, Schäden und erste Einordnung, *Tech. rep.*, Bundesamt für Umwelt BAFU, Eidgenössische Forschungsanstalt WSL, Bern.

Bhunya, P. K., R. Berndtsson, C. S. P. Ojha, and S. K. Mishra (2007), Suitability of Gamma, Chi-square, Weibull, and Beta distributions as synthetic unit hydrographs, *Journal of Hydrology*, *334*, 28–38.

Bitterli, T., P. Aviolat, R. Brändli, R. Christe, S. Fracheboud, D. Frey, M. George, F. Matousek, and J. P. Tripet (2007), Groundwater resources, in *Hydrological Atlas of Switzerland*, p. 8.6, Bern.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (2013), *Runoff prediction in ungauged basins*, 465 pp., Cambridge University Press, Cambridge.

Bormann, H., N. Pinter, and S. Elfert (2011), Hydrological signatures of flood trends on German rivers: Flood frequencies, flood heights and specific stages, *Journal of Hydrology*, *404*(1-2), 50–66, doi: 10.1016/j.jhydrol.2011.04.019.

Boughton, W., and O. Droop (2003), Continuous simulation for design flood estimation - A review, *Environmental Modelling and Software*, *18*(4), 309–318, doi: 10.1016/S1364-8152(03)00004-5.

Brázdil, R., Z. W. Kundzewicz, and G. Benito (2006), Historical hydrology for studying flood risk in Europe, *Hydrological Sciences Journal*, *51*, 739–764, doi: 10.1623/hysj.51.5.739.

Breiman, L. (2001), Random Forests, *Machine Learning*, *45*(1), 5–32, doi: 10.1023/A:1010933404324.

Brunner, M. I., O. Vannier, A.-C. Favre, D. Viviroli, P. Meylan, A. Sikorska, and J. Seibert (2016), Flood volume estimation in Switzerland using synthetic design hydrographs - a multivariate statistical approach, *Interpraevent 2016*, pp. 468–476.

Brunner, M. I., A. E. Sikorska, and J. Seibert (2017), Bivariate analysis of floods in climate impact assessments, *Science of The Total Environment*, p. in press, doi: 10.1016/j.scitotenv.2017.10.176.

Bundesamt für Statistik (2003), Geodaten der Bundesstatistik.

BWG (2003), Hochwasserabschätzung in schweizerischen Einzugsgebieten, *Tech. rep.*, Bundesamt für Wasser und Geologie BWG, Bern.

Camezind-Wildi, R. (2005), Empfehlung Raumplanung und Naturgefahren, *Tech. rep.*, Bundesamt für Raumentwicklung, Bundesamt für Wasser und Geologie, Bundesamt für Umwelt, Wald und Landschaft, Bern.

Camici, S., A. Tarpanelli, L. Brocca, F. Melone, and T. Moramarco (2011), Design soil moisture estimation by comparing continuous and storm-based rainfall-runoff modeling, *Water Resources Research*, *47*(5), doi: 10.1029/2010WR009298.

Chebana, F., and T. B. M. J. Ouarda (2008), Depth and homogeneity in regional flood frequency analysis, *Water Resources Research*, *44*(11), 1–14, doi: 10.1029/2007WR006771.

Chowdhary, H., L. A. Escobar, and V. P. Singh (2011), Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data, *Hydrology Research*, *42*, 193–216, doi: 10.2166/nh.2011.065.

Coles, S. (2001), *An introduction to statistical modeling of extreme values*, 208 pp., Springer, London.

Cuevas, A., M. Febrero, and R. Fraiman (2007), Robust estimation and classification for functional data via projection-based depth notions, *Computational Statistics*, *22*(3), 481–496, doi: 10.1007/s00180-007-0053-0.

Cullen, A. C., and H. C. Frey (1999), *Probabilistic techniques in exposure assessment*, 335 pp., Society for Risk Analysis, New York.

Dalrymple, T. (1960), Flood-Frequency Analyses, *Geological survey water supply paper 1543-A*, p. 80.

de Moel, H., B. Jongman, H. Kreibich, B. Merz, E. Penning-Rowsell, and P. J. Ward (2015), Flood risk assessments at different spatial scales, *Mitigation and Adaptation Strategies for Global Change*, *20*(6), 865–890, doi: 10.1007/s11027-015-9654-z.

Delgado, J. M., B. Merz, and H. Apel (2014), Projecting flood hazard under climate change: An alternative approach to model chains, *Natural Hazards and Earth System Sciences*, *14*(6), 1579–1589, doi: 10.5194/nhess-14-1579-2014.

Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall (2012), Merkblatt DWA-M 552, *Tech. rep.*, DWA, Hennef, Germany.

Di Baldassarre, G., F. Martinez, Z. Kalantari, and A. Viglione (2017), Drought and flood in the Anthropocene: Feedback mechanisms in reservoir operation, *Earth System Dynamics*, *8*(1), 225–233, doi: 10.5194/esd-8-225-2017.

Diezig, R., and R. Weingartner (2007), Hochwasserprozesstypen — Schlüssel zur Hochwasserabschätzung, *Wasser und Abfall*, *4*, 18–26.

Eckhardt, K. (2005), How to construct recursive digital filters for baseflow separation, *Hydrological Processes*, *19*, 507–515, doi: 10.1002/hyp.5675.

Efron, B., and R. Tibshirani (1993), *An introduction to the bootstrap*, 436 pp., Chapman & Hall/CRC, Dordrecht.

Eidgenössische Forschungsanstalt für Wald Schnee und Landschaft (WSL) (1999), *Schweizerisches Landesforstinventar. Ergebnisse der Zwietaufnahme 1993-1995*, 442 pp., BUWAL, Bern.

Evin, G., J. Blanchet, E. Paquet, F. Garavaglia, and D. Penot (2016), A regional model for extreme rainfall based on weather patterns subsampling, *Journal of Hydrology*, *541*, 1185–1198, doi: 10.1016/j.jhydrol.2016.08.024.

Frahm, G., M. Junker, and A. Szimayer (2003), Elliptical copulas: applicability and limitations, *Statistics & Probability Letters*, *63*(3), 275–286, doi: 10.1016/S0167-7152(03)00092-0.

Freund, Y., and R. R. E. Schapire (1996), Experiments with a New Boosting Algorithm, in *International conference on machine learning*, pp. 148–156, doi: 10.1.1.133.1040.

Friedman, J. H. (2001), Greedy function approximation: a gradient boosting machine, *The Annals of Statistics*, *29*(5), 1189–1232.

Friedman, J. H. (2002), Stochastic gradient boosting, *Computational Statistics & Data Analysis*, *38*(4), 367–378.

Gado, T. A., and V.-T.-V. Nguyen (2016), An at-site flood estimation method in the context of nonstationarity I. A simulation study, *Journal of Hydrology*, *535*, 710–721, doi: 10.1016/j.jhydrol.2015.12.063.

Genest, C., and A.-C. Favre (2007), Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, *12*(4), 347–367, doi: 10.1061/(ASCE)1084-0699(2007)12:4(347).

Genest, C., B. Rémillard, and D. Beaudoin (2009), Goodness-of-fit tests for copulas: A review and a power study, *Insurance: Mathematics and Economics*, *44*, 199–213, doi: 10.1016/j.insmatheco.2007.10.005.

Gordon, A. (1999), *Classification*, 2nd ed., 256 pp., Chapman & Hall/CRC, Boca Raton.

Gräler, B., M. J. van den Berg, S. Vandenberghe, A. Petroselli, S. Grimaldi, B. D. Baets, and N. E. C. Verhoest (2013), Multivariate return periods in hydrology: a critical and practical review on synthetic design hydrograph estimation, *Hydrological Earth System Sciences*, *17*, 1281–1296, doi: 10.5194/hess-17-1281-2013.

Grimaldi, S., A. Petroselli, and F. Serinaldi (2012a), Design hydrograph estimation in small and ungauged watersheds: Continuous simulation method versus event-based approach, *Hydrological Processes*, *26*(20), 3124–3134, doi: 10.1002/hyp.8384.

Grimaldi, S., A. Petroselli, and F. Nardi (2012b), A parsimonious geomorphological unit hydrograph for rainfall–runoff modelling in small ungauged basins, *Hydrological Sciences Journal*, *57*(1), 73–83, doi: 10.1080/02626667.2011.636045.

Grimaldi, S., A. Petroselli, G. Salvadori, and C. De Michele (2016), Catchment compatibility via copulas: A non-parametric study of the dependence structures of hydrological responses, *Advances in Water Resources*, *90*, 116–133, doi: 10.1016/j.advwatres.2016.02.003.

Hall, J., and D. Solomatine (2008), A framework for uncertainty analysis in flood risk management decisions, *International Journal of River Basin Management*, *6*(2), 85–98, doi: 10.1080/15715124.2008.9635339.

Hannah, D. M., B. P. G. Smith, A. M. Grunell, and G. R. McGregor (2000), An approach to hydrograph classification, *Hydrological Processes*, *14*, 317–338.

Hastie, T., R. Tibshirani, and J. Friedman (2008), *The elements of statistical learning*, Springer series in statistics, 745 pp., Springer, Stanford, California.

Heffernan, J. E. (2000), A Directory of Coeffcients of Tail Dependence, *Extremes*, *33*, 279–290.

Hofner, B., A. Mayr, N. Robinzonov, and M. Schmid (2014), Model-based boosting in R: a hands-on tutorial using the R package mboost, *Computational Statistics*, *29*(3), 3–35, doi: 10.1007/s00180-012-0382-5.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional frequency analysis*, 238 pp., Cambridge University Press, Cambridge, doi: 10.1017/CBO9780511529443.

Houghton, J. C. (1978), Birth of a parent: The Wakeby Distribution for modeling flood flows, *Water Resources Research*, *14*(6), 1105–1109, doi: 10.1029/WR014i006p01105.

Hundecha, Y., J. Parajka, and A. Viglione (2017), Flood type classification and assessment of their past changes across Europe, *Earth Syst. Sci. Discuss.*, *21*(July), doi: 10.5194/hess-2017-356.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2013), *An introduction to statistical learning. With applications in R.*, 418 pp., Springer, New York, doi: 10.1007/978-1-4614-7138-7.

Jensen, H., H. Lang, and J. Rinderknecht (1997), Extreme point rainfall of varying duration and return period 1901-1970, in *Hydrological Atlas of Switzerland*, chap. 2.4, FOEN, Bern.

Joe, H. (2014), *Dependence modeling with copulas*, 480 pp., Chapman & Hall/CRC, Boca Raton.

Klein, B., M. Pahlow, Y. Hundecha, and A. Schumann (2010), Probability analysis of hydrological loads for the design of flood control systems using copulas, *Journal of Hydrologic Engineering*, *15*(5), doi: 10.1061/(ASCE)HE.1943-5584.0000204.

Klemes, V. (1993), Probability of extreme hydrometeorological events - a different approach, *IAHS*, *213*, 167–176.

Kreibich, H., and B. Dimitrova (2010), Assessment of damages caused by different flood types, *WIT Transactions on Ecology and the Environment*, *133*, 3–11, doi: 10.2495/FRIAR100011.

Kysely, J. (2008), A Cautionary Note on the Use of Nonparametric Bootstrap for Estimating Uncertainties in Extreme-Value Models, *Journal of applied meteorology and climatology*, *47*, 3236–3251, doi: 10.1175/2008JAMC1763.1.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis (1978), Some comparisons of flood statistics in real and log space, *Water Resources Research*, *14*(5), 902–920, doi: 10.1029/WR014i005p00902.

Lang, M., T. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling, *Journal of Hydrology*, *225*, 103–117.

McGlynn, B. L., G. Blöschl, M. Borga, H. Bormann, R. Hurkmans, J. Komma, L. Nandagiri, R. Jijlenhoet, and T. Wagener (2013), A data acquisition framework for runoff prediction in ungauged basins, in *Predictions in ungauged basins. A synthesis across processes, places and scales*, edited by G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije, chap. 3, pp. 29–52, Cambridge University Press, Cambridge.

McMillan, H. K., and I. K. Westerberg (2015), Rating curve estimation under epistemic uncertainty, *Hydrological Processes*, *29*(7), 1873–1882, doi: 10.1002/hyp.10419.

Mediero, L., A. Jiménez-Alvarez, and L. Garrote (2010), Design flood hydrographs from the relationship between flood peak and volume, *Hydrology and Earth System Sciences*, *14*, 2495–2505, doi: 10.5194/hess-14-2495-2010.

Merz, B., and A. H. Thieken (2009), Flood risk curves and uncertainty bounds, *Natural Hazards*, *51*(3), 437–458, doi: 10.1007/s11069-009-9452-6.

Merz, B., et al. (2014), Floods and climate: emerging perspectives for flood risk assessment and management, *Natural Hazards and Earth System Science*, *14*(7), 1921–1942, doi: 10.5194/nhess-14-1921-2014.

Merz, R., and G. Blöschl (2003), A process typology of regional floods, *Water Resources Research*, *39*(12), 1340, doi: 10.1029/2002WR001952.

Merz, R., and G. Blöschl (2004), Regionalisation of catchment model parameters, *Journal of Hydrology*, *287*(1), 95–123, doi: 10.1016/j.jhydrol.2003.09.028.

Merz, R., and G. Blöschl (2005), Flood frequency regionalisation - spatial proximity vs. catchment attributes, *Journal of Hydrology*, *302*, 283–306, doi: 10.1016/j.jhydrol.2004.07.018.

Merz, R., and G. Blöschl (2008a), Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information, *Water Resources Research*, *44*, doi: 10.1029/2007WR006744.

Merz, R., and G. Blöschl (2008b), Flood frequency hydrology: 2. Combining data evidence, *Water Resources Research*, *44*, doi: 10.1029/2007WR006745.

Merz, R., and G. Blöschl (2009), A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, *Water Resources Research*, *45*.

MeteoSwiss (2013), Documentation of MeteoSwiss grid-data products: Daily precipitation (final analysis): RhiresD, *Tech. rep.*, MeteoSwiss.

Meyer, V., D. Haase, and S. Scheuer (2009), Flood risk assessment in European river basins–concept, methods, and challenges exemplified at the Mulde River., *Integrated environmental assessment and management*, *5*(1), 17–26, doi: 10.1897/IEAM.

Mielke, P. W. (1973), Another family of distributions for describing and analyzing precipitation data, *Journal of applied meteorology*, *12*, 275–280.

Montanari, A., and D. Koutsoyiannis (2012), A blueprint for process-based modeling of uncertain hydrological systems, *Water Resources Research*, *48*(9), 1–15, doi: 10.1029/2011WR011412.

Nadarajah, S. (2007), Probability models for unit hydrograph derivation, *Journal of Hydrology*, *344*, 185–189, doi: 10.1016/j.jhydrol.2007.07.004.

Nijssen, D., A. Schumann, M. Pahlow, and B. Klein (2009), Planning of technical flood retention measures in large river basins under consideration of imprecise probabilities of multivariate hydrological loads., *Natural Hazards and Earth System Sciences*, *9*, 1349–1363.

Ouarda, T. B. M. J., C. Girard, G. S. Cavadias, and B. Bobée (2001), Regional flood frequency estimation with canonical correlation analysis, *Journal of Hydrology*, *254*, 157–173.

Pappenberger, F., and K. J. Beven (2006), Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resources Research*, *42*(5), 1–8, doi: 10.1029/2005WR004820.

Pappenberger, F., H. Harvey, K. Beven, J. Hall, and I. Meadowcroft (2006), Decision tree for choosing an uncertainty analysis methodology: a wiki experiment, *Hydrological Processes*, *20*(17), 3793–3798, doi: 10.1002/hyp.6541.

Parajka, J., R. Merz, and G. Blöschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrology and Earth System Siences*, *9*, 157–171, doi: 10.5194/hess-9-157-2005.

Pilgrim, D. H. (1986), Bridging the Gap Between Flood Research and Design Practice, *Water Resources Research*, *22*(9), 165–176.

Pilgrim, D. H., and I. Cordery (1993), Flood runoff, in *Handbook of hydrology*, edited by D. R. Maidment, chap. 9, pp. 477–546, McGraw-Hill.

Plate, E. J. (2002), Flood risk and flood management, *Journal of Hydrology*, *267*, 2–11.

Poulin, A., D. Huard, A.-C. Favre, and S. Pugin (2007), Importance of tail dependence in bivariate frequency analysis, *Journal of Hydrologic Engineering*, *12*(4), doi: 10.1061/(ASCE)1084-0699(2007)12:4(394).

Rai, R. K., S. Sarkar, and V. P. Singh (2009), Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph, *Water Resources Management*, *23*, 899–929, doi: 10.1007/s11269-008-9306-0.

Ramsay, J. O., and B. W. Silverman (2002), *Applied functional data analysis: methods and case studies*, 190 pp., Springer, New York, doi: 10.1007/b98886.

Razavi, T., and P. Coulibaly (2013), Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods, *Journal of Hydrologic Engineering*, *18*(8), 958–975, doi: 10.1061/(ASCE)HE.1943-5584.0000690.

Requena, A. I., L. Mediero, and L. Garrote (2013), A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation, *Hydrological Earth System Sciences*, *17*, 3023–3038, doi: 10.5194/hess-17-3023-2013.

Requena, A. I., F. Chebana, and L. Mediero (2016), A complete procedure for multivariate index-flood model application, *Journal of Hydrology*, *535*, 559–580, doi: 10.1016/j.jhydrol.2016.02.004.

Rogger, M., B. Kohl, H. Pirkl, A. Viglione, J. Komma, R. Kirnbauer, R. Merz, and G. Blöschl (2012), Runoff models and flood frequency statistics for design flood estimation in Austria – Do they tell a consistent story?, *Journal of Hydrology*, *456*, 30–43, doi: 10.1016/j.jhydrol.2012.05.068.

Rosbjerg, D., et al. (2013), Prediction of floods in ungauged basins, in *Runoff prediction in ungauged basins. A synthesis across processes, places and scales*, edited by G. Blöschl, M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije, chap. 9, pp. 189–226, Cambridge University Press, Cambridge.

Salas, J. D., and J. Obeysekera (2014), Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events, *Jouarnal of hydrologic engineering. ASCE.*, *19*, 1–5, doi: 10.1061/(ASCE)HE.1943-5584.0000820.

Salinas, J. L., A. Kiss, A. Viglione, R. Viertl, and G. Blöschl (2016), A fuzzy Bayesian approach to flood frequency estimation with imprecise historical information, *Water Resources Research*, *52*, 6730–6750, doi: 10.1002/2016WR019177.

Salvadori, G. (2004), Bivariate return periods via 2-copulas, *Statistical Methodology*, *1*, 129–144.

Salvadori, G., and C. De Michele (2004), Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, *Water Resources Research*, *40*, doi: 10.1029/2004WR003133.

Salvadori, G., C. DeMichele, and F. Durante (2011), On the return period and design in a multivariate framework, *Hydrological Earth System Sciences*, *15*, 3293–3305, doi: 10.5194/hess-15-3293-2011.

Salvadori, G., G. R. Tomasicchio, and F. D'Alessandro (2014), Practical guidelines for multivariate analysis and design in coastal and off-shore engineering, *Coastal Engineering*, *88*, 1–14, doi: 10.1016/j.coastaleng.2014.01.011.

Salvadori, G., F. Durante, C. De Michele, M. Bernardi, and L. Petrella (2016), A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities, *Water Resources Research*, *52*(5), 3701–3721, doi: 10.1002/2015WR017225.

Schumann, A. H., D. Nijssen, and M. Pahlow (2010), Handling uncertainties of hydrological loads in flood retention planning, *International Journal of River Basin Management*, *8*(3-4), 281–294, doi: 10.1080/15715124.2010.512561.

Serinaldi, F. (2015), Dismissing return periods!, *Stochastic Environmental Research and Risk Assessment*, *29*, 1179–1189, doi: 10.1007/s00477-014-0916-1.

Serinaldi, F., and S. Grimaldi (2011), Synthetic design hydrographs based on distribution functions with finite support, *Journal of Hydrologic Engineering*, *16*, 434–446, doi: 10.1061/(ASCE) HE.1943-5584.0000339.

Sikorska, A. E., A. Scheidegger, K. Banasik, and J. Rieckermann (2012), Bayesian uncertainty assessment of flood predictions in ungauged urban basins for conceptual rainfall-runoff models, *Hydrology and Earth System Sciences*, *16*(4), 1221–1236, doi: 10.5194/hess-16-1221-2012.

Sikorska, A. E., D. Viviroli, and J. Seibert (2015), Flood type classification in mountainous catchments using crisp and fuzzy decision trees, *Water Resources Research*, *51*(10), 7959–7976, doi: 10.1002/2015WR017326.

Sklar, A. (1959), Fonctions de répartition à n dimensions et leurs marges, *Publ Inst Statist Univ Paris*, *8*, 229–231.

Skoien, J. O., R. Merz, and G. Blöschl (2006), Top-kriging - geostatistics on stream networks, *Hydrology and Earth System Sciences*, *10*, 277–287, doi: 10.5194/hess-10-277-2006.

Smithers, J. C. (2012), Methods for design flood estimation in South Africa, *Water SA*, *38*(4), 633–646, doi: http://dx.doi.org/10.4314/wsa.v38i4.19.

Soczynska, U., B. Nowicka, and U. Somorowska (1997), Prediction of design storms and floods, *Iahs Publication*, (246), 297–303.

Šraj, M., A. Viglione, J. Parajka, and G. Blöschl (2016), The influence of non-stationarity in extreme hydrological events on flood frequency estimation, *J. Hydrol. Hydromech.*, (November), 1–12, doi: 10.1515/johh-2016-0032.

Stedinger, J. R., and T. A. Cohn (1986), Flood Frequency Analysis With Historical and Paleoflood Information, *Water Resources Research*, *22*(5), 785–793.

Steinbakk, G. H., T. L. Thorarinsdottir, T. Reitan, L. Schlichting, S. Holleland, and K. Engeland (2016), Propagation of rating curve uncertainty in design flood estimation, *Water Resources Research*, *52*, 6897–6915, doi: 10.1002/2015WR018516.

Steinschneider, S., Y.-C. E. Yang, and C. Brown (2014), Combining regression and spatial proximity for catchment model regionalization: a comparative study, *Hydrological Sciences Journal*, *6667*(October), 1–18, doi: 10.1080/02626667.2014.899701.

Strobl, C., J. Malley, and G. Tutz (2009), An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests., *Psychological Methods*, *14*(4), 323–348, doi: 10.1037/a0016973.

Szolgay, J., L. Gaal, S. Kohnova, K. Hlavcova, R. Vyleta, T. Bacigal, and G. Blöschl (2015), A process-based analysis of the suitability of copula types for peak-volume flood relationships, in *Proceedings of the International Association of Hydrological Sciences*, pp. 183–188, IAHS, doi: 10.5194/piahs-370-183-2015.

Thieken, A., et al. (2009), Methods for the evaluation of direct and indirect flood losses, *RIMAX Contributions at the 4th International Symposium on Flood Defence (ISFD4)*, pp. 1–10.

Tung, Y.-K., and B.-C. Yen (2005), *Hydrosystems engineering uncertainty analysis*, 285 pp., McGraw-Hill Book Company, New York.

Tung, Y.-K., K.-C. Yeh, and J.-C. Yang (1997), Regionalization of unit hydrograph parameters: 1. Comparison of regression analysis techniques, *Stochastic Hydrology and Hydraulics*, *11*, 145–171.

Vandenberghe, S., N. E. C. Verhoest, E. Buyse, and B. De Baets (2010), A stochastic design rainfall generator based on copulas and mass curves, *Hydrology and Earth System Sciences*, *14*(12), 2429–2442, doi: 10.5194/hess-14-2429-2010.

Viglione, A., R. Merz, and G. Blöschl (2009), On the role of the runoff coefficient in the mapping of rainfall to flood return periods, *Hydrology and Earth System Sciences*, *6*(1), 627–665, doi: 10.5194/hessd-6-627-2009.

Viviroli, D., M. Zappa, J. Gurtz, and R. Weingartner (2009), An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools, *Environmental Modelling & Software*, *24*, 1209–1222, doi: 10.1016/j.envsoft.2009.04.001.

Wagener, T., M. Sivapalan, P. Troch, and R. Woods (2007), Catchment classification and hydrologic similarity, *Geography Compass*, *1*(4), 901–931.

Wang, Z., J. Yan, and Y. Zhang (2014), Incorporating spatial dependence in regional frequency analysis, *Water Resources Research*, *50*, 9570–9585, doi: 10.1002/2013WR014849.

Ward, J. H. (1963), Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, *58*(301), 236–244, doi: 10.1080/01621459.1963.10500845.

Wetter, O. (2017), The Potential of Historical Hydrology in Switzerland, *Hydrology and Earth System Sciences Discussions*, pp. 27–34, doi: 10.5194/hess-2017-410.

Yue, S., and P. Rasmussen (2002), Bivariate frequency analysis: discussion of some useful concepts in hydrological application, *Hydrological Processes*, *16*, 2881–2898, doi: 10.1002/hyp.1185.

Yue, S., T. Ouarda, B. Bobée, P. Legendre, and P. Bruneau (2002), Approach for describing statistical properties of flood hydrograph, *Journal of Hydrologic Engineering*, *7*(2), 147–153, doi: 10.1061/(ASCE)1084-0699(2002)7:2(147).

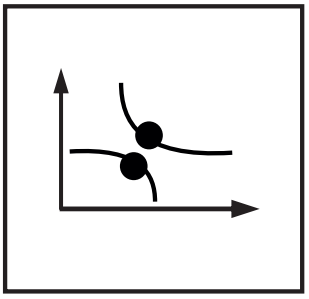
Table 1: List of stations used in this thesis. The station name is provided together with a catchment ID, catchment area [km$^2$], elevation [m.a.s.l.] (ELEV), mean elevation [m.a.s.l.] (MELEV), degree of glaciation [%] (DG), record length [a] (RL), and the owner of the station. Data owners are the Federal Office for the Environment (FOEN), and different cantons: Zürich (ZH), Vaud (VD), Solothurn (SO), Bern (BE), Baselland (BL), Aargau (AG), and Thurgau (TG).

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner |
|---|---|---|---|---|---|---|---|
| 1 | Aabach–Mönchaltorf | 46 | 440 | 521 | 0 | 34 | ZH |
| 2 | Aach–Salmsach | 49 | 406 | 480 | 0 | 40 | FOEN |
| 3 | Aire–Configon | 57 | 398 | 454 | 0 | 23 | GE |
| 4 | Allenbach–Adelboden | 29 | 1297 | 1856 | 0 | 40 | FOEN |
| 5 | Alpbach–Erstfeld | 21 | 1019 | 2200 | 28 | 41 | FOEN |
| 6 | Alp–Einsiedeln | 46 | 840 | 1155 | 0 | 23 | FOEN |
| 7 | Altbach–Bassersdorf | 13 | 470 | 549 | 0 | 37 | ZH |
| 8 | Arbogne–Avenches | 70 | 435 | 597 | 0 | 20 | VD |
| 9 | Areuse–Boudry | 377 | 444 | 1060 | 0 | 31 | FOEN |
| 10 | Arnon–Grandson | 83 | 434 | 942 | 0 | 20 | VD |
| 11 | Aubonne–Allaman | 91 | 390 | 890 | 0 | 35 | FOEN |
| 12 | Augstbach–Balsthal | 64 | 485 | 801 | 0 | 20 | SO |
| 13 | Biber–Biberbrugg | 32 | 825 | 1009 | 0 | 25 | FOEN |
| 14 | Bibere–Kerzers | 50 | 443 | 540 | 0 | 34 | FOEN |
| 15 | Birse–Court | 92 | 663 | 925 | 0 | 20 | BE |
| 16 | Birse–Moutier | 183 | 519 | 930 | 0 | 40 | FOEN |
| 17 | Birse–Soyhières | 590 | 395 | 810 | 0 | 28 | FOEN |
| 18 | Birsig–Binningen | 75 | 281 | 434 | 0 | 35 | BL |
| 19 | Birs-Münchenstein | 911 | 268 | 726 | 0 | 40 | FOEN |
| 20 | Breggia–Chiasso | 47 | 255 | 927 | 0 | 40 | FOEN |
| 21 | Brinaz–Yverdon-les-Bains | 14 | 434 | 542 | 0 | 20 | VD |
| 22 | Broye–Payerne | 392 | 441 | 710 | 0 | 40 | FOEN |
| 23 | Bruggbach–Gipf/Oberfrick | 45 | 356 | 575 | 0 | 35 | AG |
| 24 | Bünz–Muri (Hasli) | 15 | 448 | 613 | 0 | 30 | AG |
| 25 | Bünz–Othmarsingen | 111 | 390 | 533 | 0 | 37 | AG |
| 26 | Bünz–Wohlen | 53 | 421 | 555 | 0 | 34 | AG |
| 27 | Buuserbach–Maisprach | 11 | 367 | 529 | 0 | 36 | BL |
| 28 | Cassarate–Pregassona | 74 | 291 | 990 | 0 | 40 | FOEN |
| 29 | Chämtnerbach–Wetzikon | 13 | 560 | 760 | 0 | 29 | ZH |
| 30 | Chandon–Avenches | 39 | 432 | 571 | 0 | 20 | VD |
| 31 | Chli Schliere–Alpnach | 22 | 453 | 1370 | 0 | 36 | FOEN |
| 32 | Chrebsbach–St. Margarethen | 14 | 503 | 581 | 0 | 21 | TG |
| 33 | Diegterbach–Diegten | 13 | 509 | 746 | 0 | 31 | BL |
| 34 | Diegterbach–Sissach | 33 | 372 | 614 | 0 | 36 | BL |
| 35 | Dorfbach–Allschwil | 11 | 281 | 360 | 0 | 30 | BL |
| 36 | Drize–Lancy | 23 | 392 | 528 | 0 | 25 | GE |
| 37 | Dünnern–Olten | 196 | 400 | 750 | 0 | 36 | FOEN |
| 38 | Eibach–Gelterkinden | 27 | 405 | 627 | 0 | 36 | BL |
| 39 | Eibach–Zeglingen | 13 | 517 | 725 | 0 | 30 | BL |
| 40 | Emme–Eggiwil | 124 | 745 | 1189 | 0 | 39 | FOEN |
| 41 | Emme–Wiler | 939 | 458 | 860 | 0 | 40 | FOEN |
| 42 | Ergolz–Itingen | 141 | 350 | 593 | 0 | 33 | BL |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 43 | Ergolz–Liestal | 261 | 305 | 590 | 0 | 40 | FOEN |
| 44 | Ergolz–Ormalingen | 30 | 410 | 585 | 0 | 36 | BL |
| 45 | Etzgerbach–Etzgen | 25 | 308 | 478 | 0 | 34 | AG |
| 46 | Eulach–Wülflingen | 73 | 410 | 532 | 0 | 43 | ZH |
| 47 | Fisibach–Fisibach | 15 | 379 | 516 | 0 | 31 | AG |
| 48 | Flon–Oron-la-Ville | 16 | 609 | 812 | 0 | 20 | VD |
| 49 | Furtbach–Würenlos | 39 | 410 | 482 | 0 | 36 | ZH |
| 50 | Geisslibach–Furtmüli | 20 | 415 | 474 | 0 | 24 | ZH |
| 51 | Glatt–Herisau | 16 | 679 | 840 | 0 | 40 | FOEN |
| 52 | Goldach–Goldach | 50 | 399 | 833 | 0 | 23 | FOEN |
| 53 | Goneri–Oberwald | 40 | 1385 | 2377 | 14 | 23 | FOEN |
| 54 | Grenet (amont)–Pigeon | 19 | 680 | 748 | 0 | 20 | VD |
| 55 | Grossbach–Gross | 11 | 900 | 1235 | 0 | 40 | FOEN |
| 56 | Grosstalbach–Isenthal | 44 | 767 | 1820 | 9 | 40 | FOEN |
| 57 | Gürbe–Belp | 117 | 511 | 837 | 0 | 40 | FOEN |
| 58 | Gürbe–Burgistein | 54 | 568 | 1044 | 0 | 28 | FOEN |
| 59 | Haselbach–Maschwanden | 20 | 390 | 495 | 0 | 37 | ZH |
| 60 | Hintere Frenke–Bubendorf | 38 | 352 | 603 | 0 | 30 | BL |
| 61 | Hintere Frenke–Reigoldswil | 15 | 489 | 742 | 0 | 32 | BL |
| 62 | Hinterrhein–Hinterrhein | 54 | 1584 | 2360 | 17 | 35 | FOEN |
| 63 | Holzbach–Villmergen | 24 | 416 | 590 | 0 | 34 | AG |
| 64 | Homburgerbach–Thürnen | 30 | 387 | 615 | 0 | 36 | BL |
| 65 | Ilfis–Langnau | 188 | 685 | 1051 | 0 | 25 | FOEN |
| 66 | Jona–Pilgersteg | 24 | 560 | 818 | 0 | 44 | ZH |
| 67 | Jona–Rüti | 58 | 450 | 669 | 0 | 20 | ZH |
| 68 | Jonen–Zwillikon | 39 | 460 | 605 | 0 | 27 | ZH |
| 69 | Kaisterbach–Kaisten | 12 | 321 | 464 | 0 | 34 | AG |
| 70 | Kander–Hondrich | 520 | 650 | 1900 | 8 | 33 | FOEN |
| 71 | Kempt–Fehraltorf | 24 | 520 | 645 | 0 | 23 | ZH |
| 72 | Kempt–Winterthur | 60 | 450 | 588 | 0 | 33 | ZH |
| 73 | Kleine Emme–Littau | 477 | 431 | 1050 | 0 | 36 | FOEN |
| 74 | Kleine Emme–Werthenstein | 311 | 540 | 1173 | 0 | 30 | FOEN |
| 75 | Köllikerbach–Kölliken | 10 | 423 | 488 | 0 | 31 | AG |
| 76 | Langeten–Huttwil | 60 | 597 | 766 | 0 | 40 | FOEN |
| 77 | Langeten–Lotzwil | 115 | 500 | 713 | 0 | 20 | BE |
| 78 | Lonza–Blatten | 78 | 1520 | 2630 | 37 | 40 | FOEN |
| 79 | Louibach–Saanen | 62 | 1085 | 1875 | 6 | 20 | BE |
| 80 | Luthern–Nebikon | 108 | 494 | 740 | 0 | 26 | FOEN |
| 81 | Lyssbach–Lyss | 50 | 444 | 574 | 0 | 22 | BE |
| 82 | Lyssbach–Schüpfen | 23 | 505 | 616 | 0 | 22 | BE |
| 83 | Magdenerbach–Rheinfelden | 33 | 300 | 483 | 0 | 31 | AG |
| 84 | Magliasina–Magliaso | 34 | 295 | 920 | 0 | 34 | FOEN |
| 85 | Marchbach–Oberwil | 27 | 296 | 462 | 0 | 34 | BL |
| 86 | Mederbach–Marthalen | 26 | 375 | 439 | 0 | 46 | ZH |
| 87 | Mederbach–Niederwiesen | 30 | 368 | 425 | 0 | 30 | ZH |
| 88 | Mentue–Yvonand | 105 | 449 | 679 | 0 | 40 | FOEN |
| 89 | Minster–Euthal | 59 | 894 | 1351 | 0 | 40 | FOEN |
| 90 | Möhlinbach–Zeiningen | 27 | 338 | 514 | 0 | 32 | AG |
| 91 | Murg–Frauenfeld | 212 | 390 | 580 | 0 | 40 | FOEN |
| 92 | Murg–Murgenthal | 207 | 419 | 637 | 0 | 34 | FOEN |
| 93 | Murg–Wängi | 79 | 466 | 650 | 0 | 40 | FOEN |
| 94 | Näfbach–Neftenbach | 38 | 394 | 464 | 0 | 22 | ZH |
| 95 | Necker–Mogelsberg | 88 | 606 | 959 | 0 | 40 | FOEN |
| 96 | Nozon–Pré Chaillet | 45 | 440 | 882 | 0 | 21 | VD |
| 97 | Önz–Heimenhusen | 84 | 440 | 583 | 0 | 20 | BE |
| 98 | Orbe–Le Sentier | 96 | 1010 | 1210 | 0 | 21 | VD |
| 99 | Orisbach–Liestal | 21 | 315 | 515 | 0 | 33 | BL |
| 100 | Ova dal Fuorn–Zernez | 55 | 1707 | 2331 | 0 | 40 | FOEN |
| 101 | Petite Glâne–Villars-le-Grand | 85 | 433 | 560 | 0 | 20 | VD |
| 102 | Pfaffnern–Vordemwald | 39 | 417 | 517 | 0 | 34 | AG |
| 103 | Plessur–Chur | 263 | 573 | 1850 | 0 | 40 | FOEN |
| 104 | Poschiavino–La Rösa | 14 | 1860 | 2283 | 0 | 40 | FOEN |
| 105 | Promenthouse–Gland | 100 | 394 | 1037 | 0 | 28 | FOEN |
| 106 | Reppisch–Birmensdorf | 24 | 466 | 665 | 0 | 44 | ZH |
| 107 | Reppisch–Dietikon | 69 | 380 | 594 | 0 | 28 | ZH |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 108 | Riale di Pincascia–Lavertezzo | 44 | 536 | 1708 | 0 | 22 | FOEN |
| 109 | Rot–Roggwil | 54 | 436 | 586 | 0 | 25 | FOEN |
| 110 | Ruederchen–Schöftland | 19 | 463 | 614 | 0 | 34 | AG |
| 111 | Scheulte–Vicques | 73 | 463 | 785 | 0 | 22 | FOEN |
| 112 | Schmittenbach–Remigen | 13 | 385 | 523 | 0 | 32 | AG |
| 113 | Schwarzenbach–Rickenbach | 15 | 410 | 454 | 0 | 22 | ZH |
| 114 | Sellenbodenbach–Neuenkirch | 11 | 515 | 615 | 0 | 23 | FOEN |
| 115 | Sense–Thörishaus | 352 | 555 | 1068 | 0 | 36 | FOEN |
| 116 | Seymaz–Thônex | 37 | 393 | 451 | 0 | 20 | GE |
| 117 | Seyon–Valangin | 112 | 630 | 970 | 0 | 34 | FOEN |
| 118 | Simme–Oberried/Lenk | 36 | 1096 | 2370 | 35 | 40 | FOEN |
| 119 | Simme–Oberwil | 344 | 777 | 1640 | 4 | 40 | FOEN |
| 120 | Simme–Zweisimmen | 203 | 930 | 1801 | 6 | 21 | BE |
| 121 | Sinserbach–Sins | 16 | 415 | 561 | 0 | 33 | AG |
| 122 | Sionge–Vuippens | 45 | 681 | 862 | 0 | 39 | FOEN |
| 123 | Sissle–Eiken | 123 | 314 | 529 | 0 | 37 | AG |
| 124 | Sissle–Hornussen | 37 | 365 | 524 | 0 | 35 | AG |
| 125 | Somvixer Rhein–Somvix | 22 | 1490 | 2450 | 7 | 36 | FOEN |
| 126 | Sorne–Delémont | 241 | 406 | 808 | 0 | 31 | FOEN |
| 127 | Staffeleggbach–Frick | 21 | 358 | 534 | 0 | 35 | AG |
| 128 | Steinach–Steinach | 24 | 406 | 710 | 0 | 30 | FOEN |
| 129 | Steinenbach–Kaltbrunn | 19 | 451 | 1112 | 0 | 30 | FOEN |
| 130 | Stichbach–Bottighofen | 16 | 410 | 522 | 0 | 22 | TG |
| 131 | Surb–Döttingen | 67 | 335 | 511 | 0 | 34 | AG |
| 132 | Surb–Unterehrendingen | 28 | 424 | 541 | 0 | 21 | AG |
| 133 | Suze–Sonceboz | 150 | 642 | 1050 | 0 | 53 | FOEN |
| 134 | Tägerbach–Wislikofen | 14 | 390 | 551 | 0 | 32 | AG |
| 135 | Talbach–Schinznach-Dorf | 15 | 360 | 552 | 0 | 34 | AG |
| 136 | Talent–Chavornay | 66 | 440 | 670 | 0 | 20 | VD |
| 137 | Taschinasbach–Grüsch | 63 | 666 | 1768 | 0 | 34 | FOEN |
| 138 | Thur–Andelfingen | 1696 | 356 | 770 | 0 | 40 | FOEN |
| 139 | Thur–Halden | 1085 | 456 | 910 | 0 | 40 | FOEN |
| 140 | Thur–Jonschwil | 493 | 534 | 1030 | 0 | 40 | FOEN |
| 141 | Thur–Stein | 84 | 850 | 1448 | 0 | 31 | FOEN |
| 142 | Töss–Altlandenberg | 67 | 621 | 871 | 0 | 36 | ZH |
| 143 | Töss–Freienstein | 399 | 360 | 626 | 0 | 29 | ZH |
| 144 | Töss–Neftenbach | 342 | 389 | 650 | 0 | 40 | FOEN |
| 145 | Töss–Rämismühle | 127 | 524 | 790 | 0 | 35 | ZH |
| 146 | Trübbach–Räzliberg | 20 | 1430 | 2610 | 54 | 22 | FOEN |
| 147 | Ürke–Holziken | 25 | 438 | 577 | 0 | 35 | AG |
| 148 | Urnäsch–Hundwil | 65 | 746 | 1085 | 0 | 33 | FOEN |
| 149 | Vedeggio–Bioggio | 95 | 280 | 950 | 0 | 35 | FOEN |
| 150 | Venoge–Ecublens | 231 | 383 | 700 | 0 | 35 | FOEN |
| 151 | Verzasca–Lavertezzo | 186 | 490 | 1672 | 0 | 24 | FOEN |
| 152 | Veveyse–Vevey | 62 | 425 | 1108 | 0 | 30 | FOEN |
| 153 | Violenbach–Augst | 17 | 268 | 425 | 0 | 35 | BL |
| 154 | Vordere Frenke–Bubendorf | 46 | 371 | 647 | 0 | 36 | BL |
| 155 | Vordere Frenke–Waldenburg | 13 | 524 | 826 | 0 | 35 | BL |
| 156 | Weisse Lütschine–Zweilütschinen | 164 | 650 | 2170 | 18 | 40 | FOEN |
| 157 | Werrikerbach–Greifensee | 12 | 440 | 493 | 0 | 21 | ZH |
| 158 | Wigger–Zofingen | 368 | 426 | 660 | 0 | 35 | FOEN |
| 159 | Wissenbach–Boswil | 12 | 460 | 684 | 0 | 34 | AG |
| 160 | Wölflinswiler Bach–Wittnau | 17 | 395 | 600 | 0 | 30 | AG |
| 161 | Wyna–Reinach | 47 | 514 | 682 | 0 | 30 | AG |
| 162 | Wyna–Suhr | 120 | 392 | 617 | 0 | 34 | AG |
| 163 | Wyna–Unterkulm | 92 | 455 | 649 | 0 | 37 | AG |

# Paper I



Overview

# Bivariate return periods and their importance for flood peak and volume estimation

Manuela Irene Brunner,[1,2]* Jan Seibert[1,3] and Anne-Catherine Favre[2]

Estimates of flood event magnitudes with a certain return period are required for the design of hydraulic structures. While the return period is clearly defined in a univariate context, its definition is more challenging when the problem at hand requires considering the dependence between two or more variables in a multivariate framework. Several ways of defining a multivariate return period have been proposed in the literature, which all rely on different probability concepts. Definitions use the conditional probability, the joint probability, or can be based on the Kendall's distribution or survival function. In this study, we give a comprehensive overview on the tools that are available to define a return period in a multivariate context. We especially address engineers, practitioners, and people who are new to the topic and provide them with an accessible introduction to the topic. We outline the theoretical background that is needed when one is in a multivariate setting and present the reader with different definitions for a bivariate return period. Here, we focus on flood events and the different probability concepts are explained with a pedagogical, illustrative example of a flood event characterized by the two variables peak discharge and flood volume. The choice of the return period has an important effect on the magnitude of the design variable quantiles, which is illustrated with a case study in Switzerland. However, this choice is not arbitrary and depends on the problem at hand. 

How to cite this article:
*WIREs Water* 2016, 3:819–833. doi: 10.1002/wat2.1173

## INTRODUCTION

The design of hydraulic structures requires reasonable estimates for flood events that have a certain likelihood of occurrence in the catchment under consideration. These estimates are called design variables and are usually quantified for a given return period.[1] The return period is defined as the average occurrence interval which refers to the expected value of the number of realizations to be awaited before observing an event whose magnitude exceeds a defined threshold.[2,3] This definition is valid under the assumption that the phenomenon is stationary over time and each realization is independent of the previous ones.[2] The return period provides a simple, yet efficient means for risk assessment because it concentrates a large amount of information into a single number. More probable events have shorter return periods, less probable events have longer return periods.[4]

In engineering practice, the choice of the return period depends on the importance of the structure under consideration and the consequences of its

*Correspondence to: manuela.brunner@geo.uzh.ch

[1]Department of Geography, University of Zurich, Zürich, Switzerland

[2]ENSE, G-INP, LTHE, Université Grenoble–Alpes, Grenoble, France

[3]Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Conflict of interest: The authors have declared no conflicts of interest for this article.

Correction added on 21 September 2016, after first online publication: the order of the second and third authors had been interchanged.

Correction added on 21 September 2016, after first online publication: instances of underscore "_" have been changed to "∨" as the correct symbol for "OR"

failure.[5] National laws and guidelines usually fix a return period for dam design. However, they do not specify whether it refers to the peak discharge, the flood volume, or the entire hydrograph.[6] Strictly, a *T*-year hydrograph does not exist. All hydrographs are different and a frequency can only be ascribed to a particular aspect of a hydrograph, such as its peak flow, its volume, or to a particular impact such as the level of inundation.[7] However, hydrological events are not only described by one variable but by a set of correlated random variables usually consisting of the flood peak, flood volume, and duration. If more than one of these variables is significant in the design process, a univariate frequency analysis, where only one variable is considered, e.g., the peak discharge, can therefore not provide a complete assessment of the probability of occurrence of a flood event[8] and might lead to an inappropriate estimation of the risk associated with that event.[4] An overestimation of the risk is not desirable because it will increase the costs of constructing the hydraulic structure. Estimating too low design values might be even worse because it increases the risk of failure.

If two or more design variables, which are not independent from each other, are significant in the design process, one needs to consider the dependence between these variables when doing flood frequency analysis. It was shown that in such a case, a bi- or multivariate analysis where two or more variables are considered, e.g., peak discharge and flood volume, will lead to more appropriate estimates than a univariate analysis.[1,4,8] The problem of how to define a return period in a multivariate context has been addressed in several publications over the last 15 years. Several ways of defining a multivariate return period have been proposed which rely on different probability concepts. Definitions use the conditional probability, the joint probability, or can be based on the Kendall's distribution or survival function.[1,4,5,8–12] The choice of a definition for a multivariate return period is not arbitrary and depends on the problem at hand.[2]

Therefore, the goal of this study is not to present the definitive definition of a multivariate return period but to give a comprehensive overview of the tools that are available to define a return period in a multivariate context. We describe the definitions that have been proposed in previous publications, expressing them with and without copulas, and illustrate them with a practical example. This overview especially addresses engineers, practitioners, and people who are new to the topic and gives them an accessible introduction to the topic by providing the background for deciding on suitable strategies of defining a return period for a particular application. Important issues that need to be addressed when wanting to estimate design variables for a certain return period are discussed.

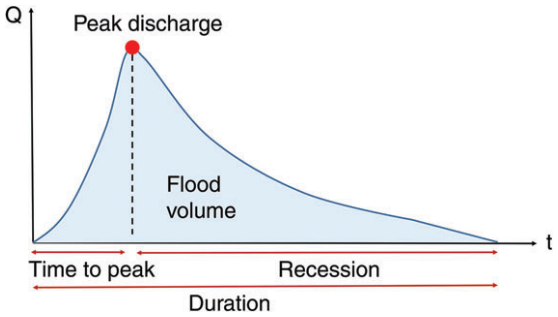



**FIGURE 1** | Illustration of different flood hydrograph characteristics.

We first provide the reader with the theoretical background that is needed when one is dealing with return periods in a multivariate setting. Then, we outline several ways of defining a bivariate return period. We provide equations only if we think that this can clarify the situation and support understanding. Following the common notation, we use upper case letters for random variables or events and lower case letters for values, parameters, or constants. Throughout this article, we use the example of a flood event characterized by the two design variables, peak discharge and flood volume, which are illustrated in Figure 1.

A third potential variable would be the flood duration, which would add a third dimension to the analysis and move us to a trivariate setting. For simplicity, we compute the volume of an event always over a window of 72 h and thereby keep the duration constant. This allows us to focus on bivariate return periods, which makes calculations less complex. However, the tools presented here are also applicable in more than two dimensions.

After a more theoretical part on return periods, the influence of the choice of a specific bivariate return period, made *a priori* according to the problem at hand, on the design variables is illustrated on a case study using data from the Birse catchment at Moutier-la-Charrue in Switzerland.

## BACKGROUND

### Practices in Estimating Design Variables

When estimating design variables for a hydraulic structure, we usually talk about design variable quantiles. The quantile can be defined as the magnitude of the event in terms of its nonexceedance probability.[13] If one considers the $p$-quantile, values in the sample have a probability of $p$% of not exceeding this quantile. The information of the nonexceedance

probability is contained in the return period. The return period is used in national guidelines to define levels of flood protection and rules for the construction of hydraulic structures. These guidelines differ from country to country but they have in common that areas and structures of lower importance are protected against events with lower return periods while inhabited areas and critical structures are protected against events with higher return periods.[14] In Switzerland, e.g., flood protection goals for agricultural land and infrastructure are based on a 20-year flood, i.e., a flood with a return period of 20 years, and protection goals for inhabited areas on a 100-year flood.[15] Very sensitive structures such as dams built for the storage of water for hydropower production have even higher protection goals. Usually, protection goals for such critical structures are based on events of a return period between 500 and 10,000 years depending on the type of the dam.[6]

## Definition of a Univariate Return Period

We need to define the univariate return period before dealing with bivariate return periods. The value of the cumulative distribution function $F_X$ of a random variable $X$ at a given value $x$ is the probability that the random variable $X$ is less than or equal to $x$

$$F_X(x) = \Pr[X \leq x]. \tag{1}$$

In hydrology, we would for example talk about the probability that the peak magnitude of a certain flood event, here denoted by $X$, is smaller than a given runoff threshold, here denoted by $x$.

In contrast, the exceedance probability that $x$ will be equaled or exceeded is given by the survival function $S_X$ of the random variable $X$, which is often used in statistical literature and stands for

$$S_X(x) = 1 - F_X(x). \tag{2}$$

If we consider our hydrological example again, we talk about the probability that the peak magnitude $X$ of a certain event exceeds a given runoff threshold $x$.

The return period $T(x)$ of the event $\{X \geq x\}$ can be written as

$$T(x) = \frac{\mu}{S_X(x)} = \frac{\mu}{1 - F_X(x)}, \tag{3}$$

where $\mu$ is the mean interarrival time between two successive events, which is defined as one divided by the number of flood occurrences per year.[8] If we look at annual maxima, $\mu$ corresponds to 1 year. In our example, $T(x)$ stands for the (univariate) return period of an event where the peak magnitude $X$ exceeds the threshold $x$.

The definition of a univariate return period can be expressed as one single equation. In practice, however, one is often faced with problems where two variables are important in the design process. For example, we often not only need to consider the flood peak, but also the flood volume. If the two variables depend on each other, we need to take into account their dependence. For this, we can look at their conditional probability of occurrence, their joint probability of occurrence or work with the Kendall's distribution or survival function. The choice of one of these probability concepts depends on the application under consideration.

Even in a bivariate context, the marginal distributions, i.e., the distributions of the single variables independent of the other variables, are of great interest. We need to analyze the marginal distributions of the design variables peak discharge and flood volume before having a look at their conditional or joint distribution.

## Marginal Distributions of Design Variables

The marginal distributions of our variables peak discharge and flood volume are linked to how we sample flood events. There are two main approaches to choose flood events from a runoff time series. The first one is the block maxima approach, which is based on choosing the highest event (usually looking at the peak discharges) over a period of time. The second approach is the peak-over-threshold (POT) approach, which is based on choosing all peaks that lie above a predefined threshold. While the block maxima approach, in which the block is defined as a year, retains only one event per year, it is possible to choose more than one event per year using the POT approach depending on the choice of the threshold.[16] After the sampling with one of these two approaches, we have a series of flood events characterized by the variables peak discharge and flood volume. Extreme value theory[17] says that block maxima follow a generalized extreme value (GEV) distribution while POT series follow a generalized Pareto distribution (GPD). The GEV model has three continuous parameters: a location parameter $\mu_l \in \mathbb{R}$, a scale parameter $\sigma > 0$, and a shape parameter $\xi \in \mathbb{R}$ and is defined as

$$F_X = \exp\left[-\left\{1 + \xi\left(\frac{x - \mu_l}{\sigma}\right)\right\}^{-\frac{1}{\xi}}\right] \quad \xi \neq 0, \tag{4}$$

defined on $\left[\xi : \left\{1 + \xi\left(\frac{x-\mu_l}{\sigma}\right)\right\} > 0\right]$.

However, the GPD uses the same parameters and is expressed as

$$F_X = 1 - \left\{1 + \xi\left(\frac{x-\mu_l}{\sigma}\right)\right\}^{-\frac{1}{\xi}} \xi \neq 0, \quad (5)$$

defined on $[x - \mu_l : \{x - \mu_l\} > 0$ and $\left\{1 + \xi\left(\frac{x-\mu_l}{\sigma}\right)\right\} > 0]$. Often, in flood frequency analysis, one works with annual maxima to guarantee the independence of the events analyzed. However, the disadvantage is that some important events are neglected because only the highest event per year is included in the data set. This problem can be solved by using a POT approach. However, even though the choice of the threshold is crucial, it is somewhat subjective.[17]

## Modeling the Dependence Between Two or More Variables

Once we defined the marginal distributions of our variables, we need to study their relationship and to assess the strength of their dependence.[18] If there is no dependence between two variables, their joint distribution is simply the product of the marginal distributions. However, if there is any dependence, we have to model their joint behavior. The cumulative distribution function $F_{XY}$ of two variables $X$ and $Y$ allows us to define the probability $F_{XY}(x, y)$ that both $X$ and $Y$ do not exceed given values $x$ and $y$[19]

$$F_{XY}(x,y) = \Pr[X \leq x, Y \leq y]. \quad (6)$$

Traditionally, the pairwise dependence between variables such as the peak, volume, and duration of flood events has been described using classical families of bivariate distributions.

The main limitation of these bivariate distributions is that the individual behavior of the two variables must be characterized by the same parametric family of univariate distributions. Copula models which are multivariate distribution functions avoid this restriction. Recent developments in statistical hydrology have shown the great potential of copulas for the construction of multivariate cumulative distribution functions and for carrying out a multivariate frequency analysis.[1,20] A list of publications on copula functions and their use in hydrology can be found on the web page of the International Commission on Statistical Hydrology[21a].

The copula approach to dependence modeling is rooted in a representation theorem due to Sklar.[22] He stated that the value of the joint cumulative distribution function $F_{XY}$ of any pair $(X, Y)$ of continuous random variables at $(x, y)$ may be written in the form of

$$F_{XY}(x,y) = C\{F_X(x), F_Y(y)\} = C(u,v), x,y \in \mathbb{R}, \quad (7)$$

where $F_X(x)$ denoted by $u$ and $F_Y(y)$ denoted by $v$ are realizations of the marginal distributions of $X$ and $Y$ whose dependence is modeled by a copula $C$. Our attention is restricted to the pair of random variables $(U, V)$, where $U$ denotes $F_X(X)$ and $V$ denotes $F_Y(Y)$. The probability integral transform allows for the conversion of the random variables $F_X(X)$ and $F_Y(Y)$ from the continuous distributions $F_X$ and $F_Y$ to the random variables $U$ and $V$ having a uniform distribution $U(0, 1)$. In our example, $F_X$ stands for the marginal distribution of the peak discharge values, $F_Y$ represents the marginal distribution of the flood volume values, and $F_{XY}$ denotes the joint distribution of peak discharges and flood volumes.

Sklar showed that $C$, $F_X$, and $F_Y$ are uniquely determined when their joint distribution $F_{XY}$ is known. The selection of an appropriate model for the dependence between $X$ and $Y$ represented by the copula can proceed independently from the choice of the marginal distributions.[23] Copulas are functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions.[24] The large number of copula families proposed in the literature allows one to choose from a large quantity of dependence structures.[23]

Five steps are involved in modeling the dependence between two or more variables with a copula:

1. Evaluation of the dependence between the variables doing an exploratory data analysis using $K$-plots and Chi-plots as well as suitable Kendall's and Spearman's independence tests.[23]
2. Choice of a number of copula families.
3. Estimation of copula parameters for each copula family.
4. Exclusion of nonadmissible copulas via suitable goodness-of-fit tests.[23,25]
5. Choice of an admissible copula via selection criteria such as the Akaike or Bayesian information criterion.[26]

For a more thorough introduction to copulas, we refer to the textbooks of Nelsen[24] or Joe[27] or the review paper by Genest and Favre.[23] For an application of copulas to estimate return periods for hydrological events, we refer to the textbook of Salvadori et al.[20]

## BIVARIATE RETURN PERIODS

A bivariate analysis is advisable when two dependent variables play a significant role in ruling the behavior of a flood.[12] In bivariate frequency analysis, in contrast to univariate frequency analysis, the definition of an event with a given return period is not unique[8]; however, it is determined by the problem at hand.[2] Salvadori et al.[5] provided a general definition of a return period which does not only apply for the univariate but also a multivariate setting and therewith helps to make the step from a univariate setting to a bivariate or multivariate framework. They defined the return period $T_D$ of a 'dangerous' event as

$$T_D = \frac{\mu}{\Pr[X \epsilon D]}, \tag{8}$$

where $D$ is a set collecting all the values judged to be dangerous according to some suitable criterion, $\mu$ is the average interarrival time of two realizations of $X$, and $\Pr[X \epsilon D]$ is the probability of a random variable (vector) $X$ to lie in the dangerous region $D$. In a setting with one significant design variable, a critical design value $x$ is used to identify the dangerous region $D$ consisting of all values exceeding $x$. In our hydrological example, $x$ would refer to a peak discharge threshold above which an event is considered dangerous. In a bivariate context, the dangerous region $D$ can be defined in various ways allowing for different return period definitions according to the problem at hand. Recently, Salvadori et al.[28] introduced the term 'hazard scenario' for a set containing all the occurrences of $X$ said to be dangerous. The ways the term return period is used in the following are all special cases of the definition given in Eq. (8).

The return period used to describe bivariate events can be determined by three types of approaches. The first of these approaches uses the conditional probability to determine a conditional return period, while the second method uses joint probability distributions to calculate joint return periods and the third approach relies on the Kendall's distribution or survival function. In hydrology, the conditional probability can for example describe the probability of a peak discharge to exceed a given threshold given that the flood volume exceeds a given threshold, or vice versa. The joint probability distributions can, e.g., describe the following two situations. First, the probability that both the peak discharge and the flood volume exceed certain thresholds during a flood event. Second, the probability that either the peak discharge or the flood volume exceed given thresholds.

The three main approaches to determine a bivariate return period are described in more detail in the next paragraphs.

### Conditional Return Period

The conditional return period approach is typically applied in situations in which one of the design variables is considered to be more important than the other one.[12] The conditional return period relies on a conditional probability distribution function of a variable given that some condition is fulfilled. The conditional return period approach can apply to particular conditional events which are chosen depending on the problem at hand. Here, we focus on two types of events that might be of special interest when designing a hydraulic structure. However, other conditional events could be investigated if necessary. The two events analyzed are described as

$$E^{>}_{X|Y} = \{X > x | \; Y > y\} \text{ and} \tag{9}$$

$$E^{>}_{Y|X} = \{Y > y | \; X > x\}, \tag{10}$$

with associated probability $\Pr[X > x | \; Y > y]$ and $\Pr[Y > y | \; X > x]$ respectively. Picking up our hydrological example again, event number one corresponds to the situation where the peak discharge $X$ exceeds a threshold $x$ given (denoted as |) that the flood volume $Y$ exceeds a threshold $y$. This event would be used if flood volume was considered to be the crucial variable. Event number two corresponds to the situation where the flood volume exceeds a threshold given that the peak discharge exceeds a predefined threshold. This event would be used if peak discharge was considered to be the most important variable in the design process.

The values of the conditional probability distribution functions for these events are defined as

$$F_{X|Y}(x,y) = 1 - \frac{F_X(x) - F_{XY}(x,y)}{1 - F_Y(y)} \text{ and} \tag{11}$$

$$F_{Y|X}(x,y) = 1 - \frac{F_Y(y) - F_{XY}(x,y)}{1 - F_X(x)}. \tag{12}$$

The conditional return period of these two conditional events can therefore be described as

$$T(x|y) = \frac{\mu}{1 - \frac{F_X(x) - F_{XY}(x,y)}{1 - F_Y(y)}} \text{ and} \tag{13}$$

$$T(y|x) = \frac{\mu}{1 - \frac{F_Y(y) - F_{XY}(x,y)}{1 - F_X(x)}}. \tag{14}$$

The conditional return period describes the mean time interval between two situations of exceedance of a certain flood volume given that a certain flood peak is exceeded or vice versa.

## Conditional Return Period Using Copulas

The study of conditional distributions can be facilitated using copulas according to Salvadori and De Michele,[4] Salvadori,[9] Renard and Lang,[29] Salvadori et al.,[20] Vandenberghe et al.,[11] Salvadori and De Michele,[10] Durante and Salvadori,[30] Salvadori et al.,[5] and Gräler et al.[1]

We consider again the two conditional events given in Eqs (9) and (10) but work with the random variables $U$ and $V$ which have a uniform distribution and stand for $F_X(X)$ and $F_Y(Y)$. Using copulas, the corresponding conditional return periods are denoted by

$$T(u|v) = \mu \frac{1-v}{1-u-v+C(u,v)} \text{ and} \quad (15)$$

$$T(v|u) = \mu \frac{1-u}{1-u-v+C(u,v)}, \quad (16)$$

where $\mu$ is the mean interarrival time between two sampled flood events.

## Joint Return Period

The joint return period of a multivariate event can be calculated using different joint probability distribution functions. Four different ways of defining values of the joint probability distribution function are illustrated in Figure 2(a). Quadrants I to IV show different ways of defining a joint probability:

Quadrant I: $\Pr[X > x, Y > y] = 1 - F_X(x) - F_Y(y) + F_{XY}(x, y) = S_{XY}(x, y)$

Quadrant II: $\Pr[X \le x, Y > y]$

Quadrant III : $\Pr[X \le x, Y \le y] = F_{XY}(x, y)$

Quadrant IV : $\Pr[X > x, Y \le y]$

In flood frequency analysis, we might either be interested in working with events situated in Quadrant I, where $X$ exceeds $x$ and $Y$ exceeds $y$, or we want to work with the events situated in Quadrants II and IV where either $Y$ exceeds $y$ or $X$ exceeds $x$.[8] These possible joint events using the OR and the AND operators, i.e., " $\vee$ " and, i.e., " $\wedge$ ", are given in Table 1.[4,9]

Continuing with our hydrological example (see Figure 2(b)), the events located in Quadrant I correspond to events where both the peak discharge $X$ and the flood volume $Y$ exceed given thresholds $x$ and $y$. Events located in Quadrant II correspond to flood events where the flood volume exceeds a given threshold but not the peak discharge. On the contrary, events located in Quadrant IV correspond to flood events where the peak discharge but not the flood volume exceeds a certain threshold.

The return period of events situated in Quadrants I, II, or IV where either peak discharge or flood volume (or both) exceeds a given threshold can be expressed by the joint OR return period (Eq. (17))

$$T^{\vee}(x,y) = \frac{\mu}{\Pr[X > x \vee Y > y]} = \frac{\mu}{1 - F_{XY}(x,y)}. \quad (17)$$

The return period of events situated in Quadrant I where both peak discharge and flood volume exceed a threshold can be expressed as the joint AND return period[6,8,31] (Eq. (18))

$$T^{\wedge}(x,y) = \frac{\mu}{\Pr[X > x \wedge Y > y]} = \frac{\mu}{1 - F_X(x) - F_Y(y) + F_{XY}(x,y)}. \quad (18)$$

## Joint Return Period Using Copulas

The bivariate joint distribution of flood peak and volume can also be obtained using a bivariate copula model.[6] Thus, the joint distribution function used for the calculation of a return period can be expressed in the form of a copula. For example, let us again consider the two events of particular interest given in Table 1, i.e.,

$$\{U > u\} \vee \{V > v\} \text{ and} \quad (19)$$

$$\{U > u\} \wedge \{V > v\}, \quad (20)$$

where $U$ stands for $F_X(X)$, the peak discharge transformed via the probability integral transform, and $V$ stands for $F_Y(Y)$, the flood volume transformed via the probability integral transform. In the first event, either the transformed peak $U$ or the transformed volume $V$ does not exceed a certain probability $u$ or $v$ respectively. In the second event, both $U$ and $V$ do not exceed a certain probability $u$ or $v$. The choice of one of these events depends, as mentioned above, on the problem at hand.

The joint OR and AND return periods of these two events using a copula can be calculated as follows

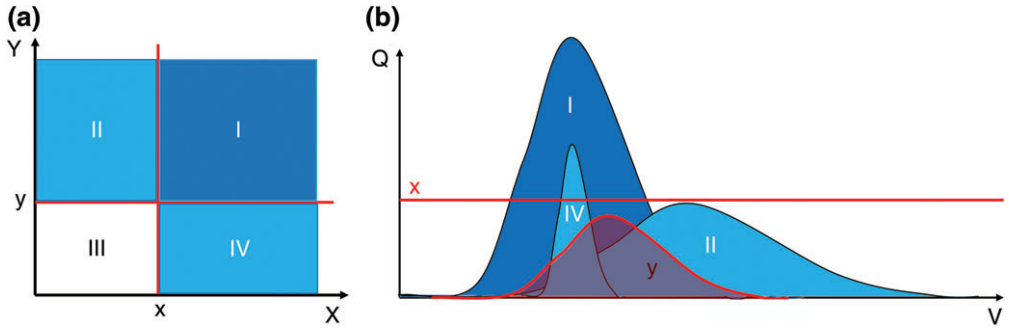



**FIGURE 2** | (a) Illustration of joint probabilities. Quadrant I shows the case when both variables *X* and *Y* exceed the values *x* and *y*. Quadrant II shows the case where *Y* but not *X* exceeds the reference value. Quadrant III shows the case where neither *X* nor *Y* exceed their reference values. Finally, Quadrant IV shows the case when *X* but not *Y* exceeds the reference value. (Modified from Ref 38; Copyright 2002) (b): Hydrological example. The red line stands for the peak discharge threshold *x*, the red hydrograph for the threshold value of total flood volume *y*. For each quadrant in figure a, one example event is given. The flood event in Quadrant I has a higher peak discharge and a higher flood volume than given by the thresholds. The event in Quadrant II has a higher volume than the threshold but a lower peak discharge. The event in Quadrant IV has a lower volume than the threshold but a higher peak discharge.

**TABLE 1** | Possible Joint Events Using the OR "∨" and the AND "∧" Operators [Correction added on 21 September 2016, after first online publication: instances of underscore "_" have been changed to "∨" as the correct symbol for "OR".]

| OR ∨ | $\{X \le x\}$ | $\{X > x\}$ |
|---|---|---|
| $\{Y \le y\}$ | $\{X \le x\} \vee \{Y \le y\}$ | $\{X > x\} \vee \{Y \le y\}$ |
| $\{Y > y\}$ | $\{X \le x\} \vee \{Y > y\}$ | $\{X > x\} \vee \{Y > y\}$ |
| AND ∧ | $\{X \le x\}$ | $\{X > x\}$ |
| $\{Y \le y\}$ | $\{X \le x\} \wedge \{Y \le y\}$ | $\{X > x\} \wedge \{Y \le y\}$ |
| $\{Y > y\}$ | $\{X \le x\} \wedge \{Y > y\}$ | $\{X > x\} \wedge \{Y > y\}$ |

Potential events of special interest in flood frequency analysis are highlighted in blue.

$$T^{\vee}(u,v) = \frac{\mu}{1 - C(u,v)} \text{ and} \tag{21}$$

$$T^{\wedge}(u,v) = \frac{\mu}{1 - u - v + C(u,v)}, \tag{22}$$

where $\mu$ is the mean interarrival time between successive events.[9] The return period only depends on the copula and not on the marginal distributions. These are just used to return from the space defined by the uniform distributions of $U$ and $V$ to the space of the real distributions of $X$ and $Y$. All pairs $(u, v)$ that are at the same probability level of the copula (i.e., they lie on an isoline of the copula) will have the same bivariate return period.

## Kendall's Return Period

Salvadori and De Michele[10] introduced the Kendall's distribution function (Eq. (23)), which depends only on the copula function $C$, and thus partitions the sample space into a supercritical and a noncritical region. The Kendall's distribution function $K_c$ stands for the cumulative distribution function of the copula's level curves or isolines and is given in a bivariate case by

$$K_c(t) = \Pr[W \le t] = \Pr[C(U,V) \le t], \tag{23}$$

where $W = C(U, V)$ is a univariate random variable.[5,10] In the bivariate case, analytical expressions for $K_c$ are available for both Archimedean and Extreme Value copulas.[20,32] When no analytical expression for $K_c$ is available, it needs to be calculated numerically based on a simulation algorithm.[5] The Kendall's distribution function allows for the calculation of the probability that a random pair $(U, V)$ in the unit square has a smaller (or larger) copula value than a given critical probability level $t$. Any critical probability level $t$ uniquely corresponds to a subdivision of the space into a supercritical and a noncritical region. The Kendall's return period therefore corresponds to the mean inter arrival time of critical events lying on the probability level $t$ which is given by

$$T_{K_c} = \frac{\mu}{1 - K_c(t)}. \tag{24}$$

Events more critical than the design event, i.e., the so-called supercritical or dangerous events, have a larger Kendall's return period than the events lying on the critical isoline and will appear much less frequently than the given design return period. This Kendall-based approach ensures that all supercritical

events have a longer joint return period than the design value, while some noncritical events might have larger marginal values than any selected design event.[1]

To overcome this issue, Salvadori et al.[33] introduced the survival Kendall's return period which yields a bounded safe region, where all the variables of interest are finite and limited. The survival Kendall's return period is based on the survival Kendall's distribution function instead of the Kendall's distribution function and is defined as

$$T_{\overline{K}_c} = \frac{\mu}{1 - \overline{K}_c(t),} \tag{25}$$

where $\overline{K}_c$ is the Kendall's survival function given by

$$\overline{K}_c(t) = \Pr[S_{XY}(X,Y) \geq t] = \Pr[C(1-U, 1-V) \geq t], \tag{26}$$

where $S_{XY}$ is the survival function of $X$ and $Y$.[12,33] The factor $1 - \overline{K}_c(t)$ yields the probability that a multivariate event will occur in the supercritical region.[12]

One of the conditional or joint return period definitions introduced above can be used to estimate design variable quantiles according to the problem at hand. However, these definitions do not take into account any interaction of the design variables peak discharge and flood volume and the hydraulic structure to be designed. To overcome this shortcoming, Volpi and Fiori[34] introduced the structure-based return period which allows for the consideration of the structure in hydraulic design in a bivariate or multivariate environment. The structure-based return period is based on the assumption that the structure design parameter $Z$ is related to the hydrological variables $X$ and $Y$ through a strictly monotonic structure function $Z = g(X, Y)$.

The return period of structure failure $T_Z$ (Eq. (27)) can be computed by applying a standard univariate frequency analysis to the random variable $Z$ using its distribution function $F_Z$:

$$T_Z = \frac{\mu}{1 - F_Z(z)}, \tag{27}$$

where $\mu$ is again the mean interarrival time between two successive events. Salvadori et al.[35] stated that it may be awkward and impractical to select the univariate law of $F_Z$ analytically, especially when the structure function is nonlinear. Therefore, they proposed to use Monte Carlo techniques to obtain an approximation of $F_Z$.

## Isolines

The difference between the univariate and the bivariate approach is that in the bivariate case, there is no unique solution of design variables associated with the return period $T$. Specific conditional or joint return periods can be achieved using various combinations of the two random variables. Hence, the bivariate return period for flood peak and flood volume must be illustrated using contour lines.[31] This return period level is a curve on a bi-dimensional graph with peak discharge and flood volume as coordinates. Based on the contours of the conditional, joint or (survival) Kendall's return periods, one can obtain various combinations of flood peaks and volumes for a given return period.[8,9] The isolines of the joint OR return period are the level curves of the copula $C$ of interest, while the isolines of the joint AND return period are the level curves of the survival copula of interest. Similarly, the conditional return period is constant over the isolines of the functions defined in Eqs (15) and (16).

## Choice of a Realization on the Isoline

In some cases of application, it might be desirable to have just one design realization or a subset of all possible realizations instead of a large set of potential events for a specified return period. Usually, one event on the isoline is chosen and declared as the design event. Several options exist for choosing one or more design realizations from the return level curve. These can be grouped into two classes. The first class of approaches aims at choosing only one design realization, whereas the second class aims at selecting a subset of design realizations on the return level curve.

Salvadori et al.[5] proposed two approaches to choose one design realization. One of these approaches looks for the 'component-wise excess design realization' whose marginal components are exceeded with the largest probability. The second approach looks for the 'most-likely design realization' taking into account the density of the multivariate distribution of the flood events. The most-likely design realization of all possible events can be obtained by selecting the point with the largest joint probability density[1,5] using Eq. (28)

$$(u,v) = \underset{C(u,v)=t}{\operatorname{argmax}} f_{XY}\left(F_X^{-1}(u), F_Y^{-1}(v)\right). \tag{28}$$

These two approaches are just two ways of choosing one design realization. In general, to identify one design realization, a suitable weight function needs to

be fixed and the point(s) where it is maximized on the critical layer can be calculated.[5,36] Salvadori et al.[12] proposed another method to choose one design realization which is applicable if one of the variables (e.g., $X$) is seen as the ruling variable (they called it $H$-conditional approach because their ruling variable was called $H$). Here, we would rather talk about the $X$- or $Y$-conditional approach. Given a return period $T$ and using a univariate approach, the corresponding critical probability level $p$ can be calculated using Eq. (3). Knowing $p$, the fitted marginal distribution of the ruling variable, $F_X$, can be inverted to provide us with a design value $x_T$ for the driving variable

$$x_T = F_X{}^{-1}(p). \tag{29}$$

Considering a particular isoline (e.g., conditional, joint, or (survival) Kendall's), a design value $y_T$ can be provided for the second variable. This corresponds to the point where the design value $x_T$ of the first variable intersects with the isoline.

The advantage of choosing just one design realization is that it is easy to handle. However, the selection of just one event reduces the amount of information that can be obtained by the multivariate approach. If one wants to keep more of this information and is rather interested in choosing a subset of design realizations from the return level curve, there are also different options. Chebana and Ouarda[13] divided the return level curve into a naïve and a proper part. The naïve part is composed of two segments starting at the end of each extremity of the proper part. The extremities are defined by the maximum values for each of the variables. An alternative is the ensemble approach proposed by Gräler et al.[1] They suggested to sample across the contours of the return level plot according to the likelihood function. By doing so, the highest density of design events is sampled around the most-likely realization, whereas less design events are sampled on the two outer limits of each contour, corresponding to the naïve part in the approach of Chebana and Ouarda.[13] Once a subset has been selected, a practitioner can still choose one design event from the subset according to the event's effect on the hydraulic structure under consideration.[37]

### Choice of Return Period Definition

The return period definition to be used in flood frequency analysis should be determined *a priori* according to the hydraulic structure to be designed or the risk assessment problem to be solved.[2] The choice of the most suitable approach to calculate the return period should be evident once the problem at hand is well defined and will affect the calculation of the design event. The different return periods do not provide answers to the same problem statement.

A univariate frequency analysis is useful when one random variable is significant in the design process. The bivariate analysis of the return periods of flood volume and flood peak may however provide more useful information for design criteria than a univariate analysis[4] if more than one variable is significant in the design process of a hydraulic structure.[2] The flood risk related to a specific event can be wrongly assessed if only the univariate return period of either the peak discharge or the flood volume is analyzed in a case where a bivariate analysis would be appropriate.[6] If two variables are significant in the design process, it is advisable to use the bivariate return period to determine the design variable quantiles. Depending on the problem at hand, one of the approaches to define a bivariate return period is chosen. The choice should be made with care and one should be aware that the approaches outlined in the previous sections provide different design variable quantiles.

The effect of the choice of one of the concepts introduced above on the design variable quantile is illustrated in the following paragraphs using a concrete example. The example shall raise the awareness of the importance of a good problem definition. As stressed by Serinaldi,[2] the choice between the possible definitions depends on how the system under consideration responds to a specific forcing. The failure mechanism of interest has a unique probabilistic description that results in a specific type of probability which in turn corresponds to a unique definition of the return period.

## EFFECT OF RETURN PERIOD CHOICE ON DESIGN VARIABLE QUANTILES

Based on the above, we can look at the effect of the choice of the return period definition, which depends on the problem at hand, on the design variable quantiles of flood peak and volume. We use flood events in a study catchment in Switzerland, the Birse at Moutier-la-Charrue, to illustrate the design variable quantiles resulting from different return period definitions. The Birse catchment lies in the Swiss Jura, has an area of 183 km$^2$, a mean elevation of 930 m. a.s.l., and no glaciers. The measurement station is situated in Moutier-la-Charrue at 519 m.a.s.l. The

measurements started in 1974 and there is a runoff time series of more than 40 years available for the analysis.

Before calculating the bivariate return period, some preparation is necessary:

1. *Sampling of flood events*. We worked with the POT approach to select four flood events per year on average from the runoff time series. This has the advantage of having all important events in the data set, even those that are not annual maxima. However, we needed to be cautious that these events were independent from each other. This was ensured by prescribing a minimum time difference between two successive events.
2. *Baseflow separation*. It is important to distinguish between the slow and the fast runoff components to analyze the statistical properties of floods.[38] We applied a recursive digital filter to separate baseflow from quick flow. Recursive digital filters have been shown to be easily applicable to a wide variety of catchments and to provide reliable results.[39]
3. *Determination of marginal distributions*. We need some information about the marginal distributions of flood peaks and volumes before we go into the bivariate analysis and consider their joint behavior. Therefore, we need to determine their marginal distributions. The events were chosen using a POT approach and therefore follow a GPD distribution according to extreme value theory.[17] Because the threshold was only applied to the peaks and not to the volumes, the volumes do not follow a GPD but a GEV distribution. The goodness-of-fit of the GPD to the peak discharges and the GEV to the flood volumes was assessed using two graphical goodness-of-fit tests, *pp*-plots and *qq*-plots, and the upper-tail Anderson Darling test proposed by Chernobai et al.[40] which showed good results. The parameters of the GEV and GPD distributions estimated for the Birse catchment are shown in Table 2.
4. *Fitting of a copula model*. The dependence between peak discharges and volumes was assessed by an exploratory data analysis using *K*-plots and Chi-plots.[23] A copula can be used to model the dependence between the two variables, peak discharge ($Q$) and flood volume ($V$). The dependence between the two variables $Q$ and $V$ was tested graphically by plotting all pairs of $Q$ and $V$ and numerically by computing two measures of dependence, Kendall's tau and Spearman's rho. Six copula models of the Archimedean copula family as well as two copulas of the meta-elliptical family were fitted using a pseudo-likelihood estimation method and tested using both graphical approaches and a goodness-of-fit test based on the Cramér-von Mises statistic.[23,25] A *p*-value for the Cramér-von Mises statistic of each copula was estimated using a bootstrap procedure.[25] The copulas which were not rejected at a level of significance of 0.05 in most catchments were found to be the Joe and the Gumbel copula. We decided to work with the Joe copula because it was rejected in fewer catchments than the Gumbel copula. The Joe copula is described by

$$C(u,v) = 1 - \left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right]^{\frac{1}{\theta}}. \quad (30)$$

It takes $\theta$ parameters in $[1\ \infty)$[24] and is able to model the dependence in the data. The parameter $\theta$ for the Birse catchment at Moutier-la-Charrue was estimated to be 1.92.

Knowing the copula found to model the dependence between peak discharges and flood volumes well, we can calculate the bivariate return period chosen to be suitable for the analysis. Table 3 shows the design quantiles for a return period of 100 years obtained for the Birse catchment by applying the different approaches outlined above. The results of all approaches are visualized in Figure 3.

**TABLE 2** | Parameters of the GEV and GPD Distributions Estimated for the Birse Catchment at Moutier-la-Charrue

| Parameter/Distribution | GEV | GPD |
|---|---|---|
| Location ($\mu$) | 294.1 | 12.9 |
| Scale ($\sigma$) | 93.1 | 11.6 |
| Shape ($\xi$) | 0.21 | −0.15 |

GEV, generalized extreme value; GPD, generalized Pareto distribution.

## Conditional Approach

We chose two different types of conditional events to illustrate the conditional approach (Eqs (9) and (10)). However, if desired, one could work with different types of conditional events such as $\{X > x \mid Y < y\}$ or $\{Y > y \mid X < x\}$. If the flood volume is considered to be the most significant variable for the design

**TABLE 3** | Design Variable Quantiles for the Peak Discharge and Flood Volume in the Birse Catchment at Moutier-la-Charrue for the Following Approaches for a Return Period of 100 Years: Univariate, $Q_{max}|V$, $V|Q_{max}$, $Q_{max}$-conditional, $V$-conditional, Joint AND, joint or, Kendall's, and Survival Kendall's

| Approach | Univariate | $Q_{max}\|V$ | $V\|Q_{max}$ | $Q_{max}$-conditional | $V$-conditional | Joint AND | Joint OR | Kendall's | Survival Kendall's |
|---|---|---|---|---|---|---|---|---|---|
| Quantiles $Q$ [$m^3/s$] | 58.8 | 60.2 | 13.1 | 58.9 | 65.1 | 55.8 | 60.6 | 54.2 | 57.4 |
| Quantiles $V$ [$Mm^3$] | 4.2 | 0.8 | 4.2 | 5.5 | 4.2 | 3.7 | 4.5 | 3.5 | 3.9 |

If the approach provides us with an isoline, the most probable event on that isoline was chosen.

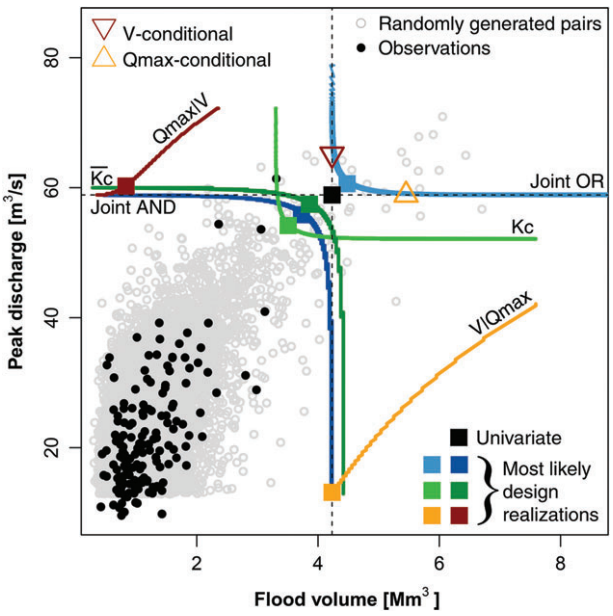



**FIGURE 3** | Design variable quantiles for different return period definitions. The black dots stand for the observed flood events for the Birse at Moutier-la-Charrue and the gray dots are 10,000 randomly generated pairs using the bivariate distribution of the peak discharges and flood volumes. The black square stands for the univariate quantile. The triangles represent the design variable pairs resulting from the $Q_{max}$- and $V$-conditional approaches applied to the joint OR isoline. The isolines represent the return level curves for the two joint approaches AND and OR, and the approaches using the Kendall's and survival Kendall's distribution function. The squares on the isolines stand for the most-likely design realizations on these isolines.

process, we work with the event given in Eq. (9) and call the approach $Q_{max}|V$. On the contrary, if the peak discharge is considered to be most important, we work with the event given in Eq. (10) and call the approach $V|Q_{max}$.

The design quantiles using these conditional approaches were calculated using Eqs (15) and (16). We retained the pairs $(u, v)$ that were located along the probability level $t$ corresponding to the given return period $T$ such that $1-t=\frac{1-u-v+C(u,v)}{1-v}$ when looking at the first event and $1-t=\frac{1-u-v+C(u,v)}{1-u}$ when looking at the second event.

All the pairs of probabilities $(u, v)$ that are at the same probability level $t$ are eligible because they correspond to the return period $T$. The design variable pairs were then calculated according to their marginal distributions $F_Q$ and $F_V$

$$Q_T = F_Q{}^{-1}(u) \text{ and} \tag{31}$$

$$V_T = F_V{}^{-1}(v). \tag{32}$$

Therefore, in contrast to the univariate case, there is no unique solution of the design variables associated with the return period $T$. Instead, all the possible solutions are located along the return period level, which is a curve on a bi-dimensional graph with $Q$ and $V$ as coordinates.

Figure 3 shows the conditional return period levels for the two conditional approaches discussed above ($Q_{max}|V$ and $V|Q_{max}$). If desired, one design variable pair on the isoline can be selected, e.g., by choosing the most probable design realization (see Table 3 and squares on isoline in Figure 3).

## Joint Approaches

If the problem at hand requires a joint analysis of peak discharges and volumes, a joint approach is appropriate. The joint approach does not provide a single design quantile pair for a given return period, but a set of pairs, all having the same return period. As mentioned in the theoretical part, two possible approaches to compute a joint return period are the AND and the OR approach.

The design quantiles using the OR approach were calculated using Eq. (21). Equation (22) was used in the AND approach. We retained the pairs $(u, v)$ that were located along the probability level $t$ corresponding to the given return period $T$ such that $1 - t = 1 - C(u, v)$ in the OR approach, and $1 - t = 1 - u - v + C(u, v)$ in the AND approach.

All the pairs of probabilities $(u, v)$ that are at the same probability level $t$ are eligible because they correspond to the return period $T$. The design variable pairs were then calculated according to their marginal distributions $F_Q$ and $F_V$ using Eqs (31) and (32).

Therefore, in contrast to the univariate case, there is no unique solution of the design variables associated with the return period $T$. Instead, all the possible solutions are located along the return period level, which is a curve on a bi-dimensional graph with $Q$ and $V$ as coordinates.

Figure 3 shows the joint return period levels for the AND and OR approaches. While the joint return period level using the AND approach is concave, the joint return period level using the OR approach is convex. Generally, the AND approach provides lower design variable quantiles than the OR approach for a given return period. If desired, one design variable pair on the isoline can be selected, e.g., by choosing the most-likely design realization (see Table 3 and squares on isoline in Figure 3) or by applying the $H$-conditional approach (here, the $Q_{max}$-conditional and $V$-conditional approaches) proposed by Salvadori et al.[12]

## Kendall's Approach

The design quantiles using the Kendall's approach were calculated using Eq. (24). The Kendall's quantile $q_t$ for the probability level $t$ could then be computed as

$$q_t = K_c^{-1}(t), \tag{33}$$

where $K_c^{-1}$ is the inverse of the Kendall's distribution $K_c$. We estimated $q_t$ using a bootstrap technique.[5] We retained the pairs $(u, v)$ that are located along the critical probability level $q_t$. All the pairs of probabilities $(u, v)$ that are at the same probability level $q_t$ are eligible because they correspond to the return period $T$. The design variable pairs were then calculated according to their marginal distributions $F_Q$ and $F_V$ using Eqs (31) and (32).

The isoline corresponding to the critical events according to the Kendall's return period is displayed in Figure 3. The event with the highest likelihood on this isoline is also indicated and given in Table 3.

Similary to the Kendall's distribution function, the survival Kendall's distribution function can be used to derive a survival Kendall's quantile instead of the Kendall's quantile

$$q_t = \overline{K}_c^{-1}(t), \tag{34}$$

where $\overline{K}_c^{-1}$ is the inverse of the Kendall's survival function.[12,33] The isoline corresponding to the critical events according to the survival Kendall's return period is also displayed in Figure 3. The event with the highest likelihood on this isoline is indicated with a square and given in Table 3.

The results presented above for different approaches to compute design variable quantiles demonstrate that the choice of the approach has a significant influence on the outcome of the design variable quantiles and that it is therefore essential to well define the problem at hand to make a suitable choice of a return period definition. Compared to the univariate quantile, the choice of the joint OR approach resulted in higher design variable quantiles. In contrast, the choice of the conditional approaches, the joint AND approach, the Kendall's approach and the survival Kendall's approach resulted in lower design variable quantiles than in the univariate case. Serinaldi[2] emphasized that this choice is not arbitrary and depends on the problem at hand. If not only the problem at hand but also the interaction of the design variables $X$ and $Y$ with the structure under consideration is well defined, the structure-based return period introduced by Volpi and Fiori[34] can be applied to derive the design variable quantile

$$z_T = F_Z^{-1}\left(1 - \frac{\mu}{T}\right), \tag{35}$$

where $F_Z^{-1}$ is the inverse of the distribution function of the design parameter $Z$. In practice, the structure function $g(X, Y)$ relating the hydrological variables peak discharge and volume to the design parameter $Z$ might be quite complex. For a specific example of a structure function, we refer to Volpi and Fiori[34] and to Salvadori et al.[35]

## Choice of a Point on the Isoline

It was mentioned earlier on that there are several possibilities to choose one design realization on the return level curve or isoline. These possibilities are illustrated in Figure 4. The isoline can be divided into a central and a naïve part. The determination of the component-wise excess realization or the most-likely design realization is a way of choosing just one realization on the critical level. If one would not like to go that far, a possibility is to work with an ensemble sampled according to the probability distribution.

## Uncertainty of Design Variable Quantiles

Independent of the return period definition used to estimate the design variable quantiles, the design variable

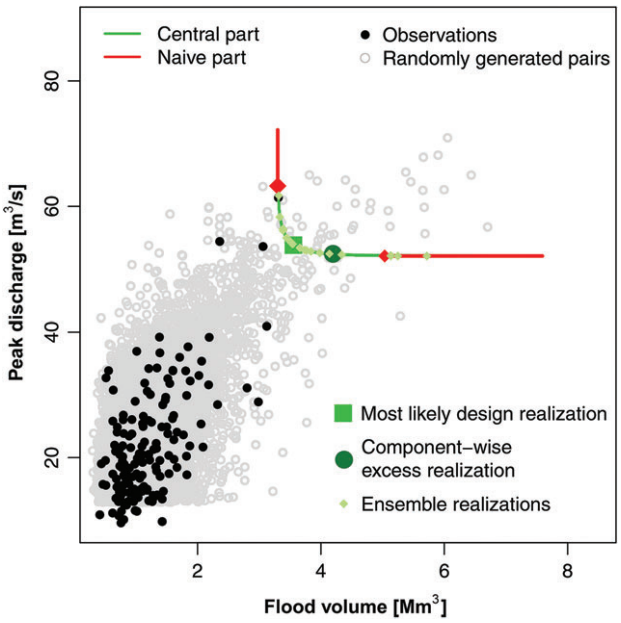



**FIGURE 4** | Kendall's critical level divided into a central (green) and two naïve parts (red). The black dots stand for the observed flood events for the Birse at Moutier-la-Charrue and the gray dots are 10,000 randomly generated pairs using the bivariate distribution of the peak discharges and flood volumes. The two possibilities of choosing one design realization are displayed. Namely, these are the most-likely design realization and the component-wise excess realization. It is also shown how a subset of realizations can be chosen with the ensemble approach.

quantiles have to be complemented with information about their uncertainty.[41] Estimated design variable quantiles are uncertain because they are made for events whose frequency goes beyond the range that is supported by the length of the flood records.[42] In a bivariate framework, the uncertainty related to the limited sample size and the uncertainty of the marginal distributions combine with the uncertainty of the dependence structure between the two variables.[18] The scientific community agrees that the uncertainty stemming from flood frequency analysis should be properly acknowledged not only in a univariate but also in a multivariate setting.[12,18,35,43,44] Serinaldi[43] recommended communicating the results of hydrological frequency analysis by complementing accurate point estimates with realistic confidence intervals. The uncertainty of extreme quantiles can be assessed using bootstrapping methods not only in univariate but also in multivariate frequency analysis.[43] Serinaldi[43] and Dung[44] proposed nonparametric and parametric bootstrap algorithms to construct confidence intervals for design variable quantiles. Practical procedures for assessing the uncertainty of multivariate design occurrences via bootstrap approaches have recently been outlined in Salvadori et al.[28]

Serinaldi[41] distinguished between three different types of uncertainty in a statistical analysis: natural uncertainty which represents the randomness and complexity of the natural process; statistical uncertainty which is related to the estimation of the parameters; and model uncertainty which depends on the choice of the statistical model. While natural uncertainty can not be reduced, statistical uncertainty can be reduced by increasing the sample size and model uncertainty by increasing the knowledge of the process under study. Serinaldi[43] and Dung[44] stated that the major source of uncertainty in the estimation of design variable quantiles is the limited sample size while parameter estimation methods and model selection are of only minor importance. When the sample is small, many joint distributions and copulas can fit the data adequately and goodness-of-fit tests cannot discriminate between alternative models because of the lack of power. Very large samples are needed to reliably estimate the marginal and joint extreme quantiles.[43] The uncertainty of design variable quantile estimates can be reduced by information expansion such as the inclusion of documentary records of historical floods or data pooling from similar catchments.[42,44] Uncertainty can also be reduced using Bayesian techniques allowing for the incorporation of different sources of information.[45]

## DISCUSSION AND CONCLUSIONS

The results presented above for different approaches to compute design variable quantiles demonstrated that the choice of the approach has a significant influence on the outcome of the design variable quantiles. The case study for a catchment in Switzerland showed that a univariate analysis can not provide a complete assessment of the probability of occurrence of a flood event if two or more dependent variables are significant in the design process. This confirms earlier results,[6,8] that univariate approaches might lead to an inadequate estimation of the risk associated with a given event. Given a specific problem, a solution to the problem of how to define a multivariate return period can be found.[2] The approaches of defining a return period discussed in this review, resulted in different design event estimates. This implies that addressing the question of how to specify the engineering problem and therewith to define a bivariate return period is important. It is impossible to provide the reader with a general suggestion for an approach to estimate multivariate design events since that depends on the problem he or she is facing. Still, this study should give him or her an overview on the

methods involved in defining a return period once he or she has outlined the problem. This study provides a basis for the practitioner or engineer to decide which of the strategies to define a return period is most suitable in his or her case. In particular, the theoretical background of five different approaches to compute design variable quantiles using conditional and joint probabilities is described, and the challenge of defining a return period was discussed with respect to flood events looking at the two variables peak discharge and flood volume. However, the analysis is neither restricted to floods nor to two variables. The concepts discussed above are also applicable in a context where more than two dependent variables are important and in other areas of application. Though, recently, it has even been questioned whether the return period and the corresponding design quantiles do actually matter in system design and planning. Serinaldi[2] strongly recommended assessing the risk of failure instead, which is the probability to observe a critical event at least once in $M$ years of the design life of a structure. The risk of failure has a unique definition independent of the nature of data and allows the consideration of both independent and dependent variables in stationary but also nonstationary settings. A multivariate failure approach to assess hydrological risk in a general and consistent mathematical way seems valuable and has recently been outlined by Salvadori et al.[28]

## NOTE

[a] http://www.stahy.org/Topics/CopulaFunction/tabid/67/Default.aspx

## ACKNOWLEDGMENTS

We thank the Federal Office for the Environment (FOEN) for financing the project via the contract 13.0028. KP/M285-0623 and for providing runoff measurement data. Detailed comments of Francesco Serinaldi and an anonymous reviewer helped to improve this article.

## REFERENCES

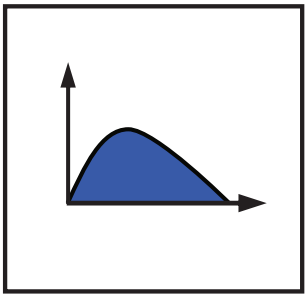
1. Gräler B, van den Berg MJ, Vandenberghe S, Petroselli A, Grimaldi S, Baets BD, Verhoest NEC. Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrol Earth Syst Sci* 2013, 17:1281–1296.
2. Serinaldi F. Dismissing return periods!. *Stoch Environ Res Risk Assess* 2015, 29:1179–1189. doi:10.1007/s00477-014-0916-1.
3. Volpi E, Fiori A, Grimaldi S, Lombardo F, Koutsoyiannis D. One hundred years of return period: strengths and limitations. *Water Resour Res* 2015, 51:8570–8585. doi:10.1002/2015WR017820.
4. Salvadori G, and De Michele C. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resour Res* 2004, 40:W12511. doi:10.1029/2004WR003133.
5. Salvadori G, DeMichele C, Durante F. On the return period and design in a multivariate framework. *Hydrol Earth Syst Sci* 2011, 15:3293–3305.
6. Requena AI, Mediero L, Garrote L. A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation. *Hydrol Earth Syst Sci* 2013, 17:3023–3038.
7. Robson A, Reed D. *Statistical Procedures for Flood Frequency Estimation: Volume 3 of the Flood Estimation Handbook*. Wallingford: Centre for Ecology & Hydrology; 2008, 338.
8. Yue S, Rasmussen P. Bivariate frequency analysis: discussion of some useful concepts in hydrological application. *Hydrol Process* 2002, 16:2881–2898.
9. Salvadori G. Bivariate return periods via 2-copulas. *Stat Methodol* 2004, 1:129–144.
10. Salvadori G, and De Michele C. Multivariate multiparameter extreme value models and return periods: A copula approach. *Water Resour Res* 2010, 46: W10501. doi:10.1029/2009WR009040.
11. Vandenberghe S, Verhoest NEC, Onof C, and De Baets B. A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall, *Water Resour Res* 2011, 47:W07529. doi:10.1029/2009WR008388.
12. Salvadori G, Tomasicchio GR, D'Alessandro F. Practical guidelines for multivariate analysis and design in coastal and off-shore engineering. *Coast Eng* 2014, 88:1–14.
13. Chebana F, Ouarda TBMJ. Multivariate quantiles in hydrological frequency analysis. *Environmetrics* 2009, 22:63–78.
14. Hingray B, Picouet C, Musy C. *Hydrology. A Science for Engineers*. Lausanne: CRC Press; 2014, 592.

15. Weingartner R. *Regionalhydrologische Analysen—Grundlagen und Anwendungen*. Beiträge zur Hydrologie der Schweiz, vol. 37. Bern: Schweizerische Gesellschaft für Hydrologie und Limnologie (SGHL); 1999, 178.
16. Lang M, Ouarda TBMJ, Bobée B. Towards operational guidelines for over-threshold modeling. *J Hydrol* 1999, 225:103–117.
17. Coles S. *An introduction to statistical modeling of extreme values*. London: Springer; 2001, 208.
18. Serinaldi F. Can we tell more than we can know? The limits of bivariate drought analyses in the United States. *Stoch Environ Res Risk Assess* 2015, 30:1691–1704. doi:10.1007/s00477-015-1124-3.
19. Kotz S, Balakrishnan N, Johnson NL. *Continuous Multivariate Distributions, Volume 1, Models and Applications*. 2nd ed. New York: Wiley; 2000, 752.
20. Salvadori G, De Michele C, Kottegoda NT, Rosso R. *Extremes in Nature. An Approach Using Copulas*, vol. 56. Dordrecht: Springer; 2007, 292.
21. International Commission on Statistical Hydrology. Copula function. Available at: http://www.stahy.org. (Accessed April 20, 2016)
22. Sklar A. Fonctions de répartition à n dimensions et leurs marges. *Publ Inst Statist Univ Paris* 1959, 8:229–231.
23. Genest C, Favre A-C. Everything you always wanted to know about copula modeling but were afraid to ask. *J Hydrol Eng* 2007, 12:347–367.
24. Nelsen RB. *An Introduction to Copulas*. 2nd ed. New York: Springer Science & Business Media; 2005, 269.
25. Genest C, Rémillard B, Beaudoin D. Goodness-of-fit tests for copulas: a review and a power study. *Insur Math Econ* 2009, 44:199–213.
26. Meylan P, Favre A-C, Musy A. *Predictive Hydrology: A Frequency Analysis Approach*. St. Helier: Science Publishers; 2012, 212.
27. Joe H. *Dependence modeling with copulas*. Vancouver: Chapman & Hall/CRC; 2014, 459.
28. Salvadori G, Durante F, Michele CD, Bernardi M, Petrella L. A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. *Water Resour Res* 2016, 52:3701–3721. doi:10.1002/2015WR017225.
29. Renard B, Lang M. Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Adv Water Resour* 2007, 30:897–912.
30. Durante F, Salvadori G. On the construction of multivariate extreme value models via copulas. *Environmetrics* 2010, 21:143–161.
31. Shiau JT. Return period of bivariate distributed extreme hydrological events. *Stoch Environ Res Risk Assess* 2003, 17:42–57.
32. Ghoudi K, Khoudraji A, Rivest EL-P. Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Can J Stat* 1998, 26:187–197. doi:10.2307/3315683.
33. Salvadori G, Durante F, Michele CD. Multivariate return period calculation via survival functions. *Water Resour Res* 2013, 2013:2308–2311. doi:10.1002/wrcr.20204.
34. Volpi E, Fiori A. Hydraulic structures subject to bivariate hydrological loads: return period, design, and risk assessment. *Water Resour Res* 2014, 50:885–897. doi:10.1002/2013WR014214.
35. Salvadori G, Durante F, Tomasicchio GR, D'Alessandro F. Practical guidelines for the multivariate assessment of the structural risk in coastal and offshore engineering. *Coast Eng* 2015, 95:77–83.
36. Salvadori G, Michele CD. Chapter 5: multivariate extreme value methods. In: AghaKouchak A, Easterling D, Hsu K, Schubert S, Sorooshian S, eds. *Extremes in a Changing Climate: Detection, Analysis and Uncertainty*. Dordrecht: Springer Science+Business Media; 2013, 115–162.
37. Volpi E, Fiori A. Design event selection in bivariate hydrological frequency analysis. *Hydrol Sci J* 2012, 57:1506–1515. doi:10.1080/02626667.2010.726357.
38. Yue S, Ouarda TBMJ, Bobée B, Legendre P, Bruneau P. Approach for describing statistical properties of flood hydrograph. *J Hydrol Eng* 2002, 7:147–153.
39. Eckhardt K. How to construct recursive digital filters for baseflow separation. *Hydrol Process* 2005, 19:507–515.
40. Chernobai A, Rachev ST, Fabozzi FJ. Composite goodness-of-fit tests for left-truncated loss samples. In: Lee C-F, Lee J, eds. *Handbook of Financial Econometrics and Statistics*. New York: Springer Science+Business Media; 2015, 575–596.
41. Serinaldi F. Assessing the applicability of fractional order statistics for computing confidence intervals for extreme quantiles. *J Hydrol* 2009, 376:528–541.
42. Reed DW. Reinforcing flood-risk estimation. *Philos Trans R Soc Lond B Biol Sci* 2002, 360:1373–1387.
43. Serinaldi F. An uncertain journey around the tails of multivariate hydrological distributions. *Water Resour Res* 2013, 2013:6527–6547. doi:10.1002/wrcr.20531.
44. Dung NV. Handling uncertainty in bivariate quantile estimation—an application to flood hazard analysis in the Mekong Delta. *J Hydrol* 2015, 527:704–717. doi:10.1016/j.jhydrol.2015.05.033.
45. Zhang Q, Xiao M, Singh VP. Uncertainty evaluation of copula analysis of hydrological droughts in the East River basin, China. *Glob Planet Change* 2015, 129:1–9. doi:10.1016/j.gloplacha.2015.03.001.

# Paper II



AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019535

# Flood type specific construction of synthetic design hydrographs

**Manuela I. Brunner[1,2], Daniel Viviroli[1,3], Anna E. Sikorska[1,4], Olivier Vannier[5], Anne-Catherine Favre[2], and Jan Seibert[1,6]**

[1]Department of Geography, University of Zurich, Zurich, Switzerland, [2]Université Grenoble-Alpes, Grenoble INP, Grenoble, France, [3]belop GmbH, Sarnen, Switzerland, [4]Department of Hydraulic Engineering, Warsaw University of Life Sciences, Warsaw, Poland, [5]Compagnie Nationale du Rhône, Lyon, France, [6]Department of Earth Sciences, Uppsala University, Uppsala, Sweden

**Key Points:**
- Construction of synthetic design hydrographs using frequency and process-based knowledge
- Flood types have a specific behavior regarding event shapes and the dependence between flood peaks and volumes
- Flood type specific design hydrographs can aid flood-risk management

**Correspondence to:**
M. I. Brunner,
manuela.brunner@geo.uzh.ch

**Citation:**
Brunner, M. I., D. Viviroli, A. E. Sikorska, O. Vannier, A.-C. Favre, and J. Seibert (2017), Flood type specific construction of synthetic design hydrographs, *Water Resour. Res.*, *53*, doi:10.1002/2016WR019535.

Received 20 JUL 2016
Accepted 9 DEC 2016
Accepted article online 22 DEC 2016

**Abstract** Accurate estimates of flood peaks, corresponding volumes, and hydrographs are required to design safe and cost-effective hydraulic structures. In this paper, we propose a statistical approach for the estimation of the design variables peak and volume by constructing synthetic design hydrographs for different flood types such as flash-floods, short-rain floods, long-rain floods, and rain-on-snow floods. Our approach relies on the fitting of probability density functions to observed flood hydrographs of a certain flood type and accounts for the dependence between peak discharge and flood volume. It makes use of the statistical information contained in the data and retains the process information of the flood type. The method was tested based on data from 39 mesoscale catchments in Switzerland and provides catchment specific and flood type specific synthetic design hydrographs for all of these catchments. We demonstrate that flood type specific synthetic design hydrographs are meaningful in flood-risk management when combined with knowledge on the seasonality and the frequency of different flood types.

## 1. Introduction

Accurate flood estimates are needed for the design of hydraulic structures and for flood-risk management. The major quantity of interest in flood estimation is the magnitude of the flood peak corresponding to a specific return period [*Rosbjerg et al*., 2013]. Flood peaks, however, provide only a limited description of a flood event. For the prevention of flood damage and for designing hydraulic structures, it is also important to know the flood volume and the shape of the flood hydrograph [*Mediero et al*., 2010]. Design flood hydrographs provide this information for any specified return period and, hence, unite the physical properties of a flood event and statistical information about the event rarity [*Serinaldi and Grimaldi*, 2011]. Design flood hydrographs are hydrographs of a suitable probability and magnitude adopted to ensure the safety of hydraulic structures [*Xiao et al*., 2009], such as dam spillways, bridges, road culverts, levees, or retention basins. Such hydrographs are also used to estimate sediment loads [*Rickenmann*, 1997], and to draw hazard maps for land use and urban planning. Design flood hydrographs contain information on several dependent design variables such as peak magnitude, flood volume, and duration, which together determine the severity of a flood. Different flood hydrograph shapes may cause differences in the costs of hydraulic structures and influence flood-control policies and flood management strategies [*Yue et al*., 2002].

Two approaches are used for design flood estimation, probabilistic and deterministic methods [*Smithers*, 2012; *Rogger et al*., 2012]. Although these are fundamentally different, they are often confused [*Pilgrim and Cordery*, 1993]. While probabilistic methods are based on the analysis of relatively long flood records, deterministic methods are based on rainfall data and take into account some catchment processes. The latter are based on the critical assumption [*Viglione et al*., 2009] that the return period of rainfall and resulting discharge are the same and require the choice of an antecedent soil moisture input [*Pilgrim and Cordery*, 1993; *Rogger et al*., 2012]. A combination of the two approaches involves the calibration of a rainfall-runoff model [*Boughton and Droop*, 2003; *Pathiraja et al*., 2012; *Pilgrim and Cordery*, 1993], running it with stochastically generated rainfall to obtain a long flood record, and a subsequent probabilistic analysis [*Rogger et al*., 2012]. For the design of minor works, such as bridges and culverts, farm dam spillways, and urban drainage systems, where a certain risk of failure is acceptable, a simple method is needed that is easily applicable by

designers with little hydrologic expertise, is physically sound, gives reproducible answers [*Pilgrim and Cordery*, 1993], and does not involve the calibration of a continuous simulation model [*Boughton and Droop*, 2003].

In this study, we propose a simple probabilistic method that is based on observed runoff data only and thus does not require the use of any rainfall-runoff model. The method estimates peak magnitude and flood volume using frequency analysis and uses a prescribed mathematical function to model the flood shape. The combination of the flood variable estimates and the flood shape results in a *synthetic* design hydrograph (SDH).

Several methods have been proposed to derive unit hydrographs that can serve as the basis for design flood hydrographs. *Yue et al*. [2002] gave an overview on existing unit hydrograph methods and grouped them into four types: traditional unit hydrograph (TUH), synthetic unit hydrograph (SUH), typical hydrograph (TH), and statistical methods (SM). The TUH of a watershed is defined as the direct runoff hydrograph resulting from a unit volume of excess rainfall of constant intensity and uniformly distributed over the drainage area [*Ramirez*, 2000]. The SUH is defined on the basis of catchment characteristics or expert knowledge and often assumes a triangular shape of the hydrograph, whereas the TH selects a typical flood hydrograph (usually the flood hydrograph with the highest peak discharge or flood volume) from sampled flood series [*Yue et al*., 2002]. *Bhunya et al*. [2007] stated that the SM, which are based on probability density functions (PDFs), are more suitable to derive unit hydrographs than traditional methods because the area under the curve is guaranteed to be equal to one and can therefore be used as the basis for a design flood hydrograph. Further, probability density functions are quite flexible and can take various shapes. *Yue et al*. [2002] proposed a method that employs the PDF of the Beta distribution, which allows a unit hydrograph to be represented with different shape types, due to the flexibility of the Beta distribution, that can be used as the basis for a design flood hydrograph. *Nadarajah* [2007] used nine additional PDFs to derive the basis for a design flood hydrograph, namely, the Lognormal, inverse Gamma, Kumaraswamy, Two-sided-power, Pareto, inverse Gaussian, F, Weibull, and Fréchet density functions. He provided estimates of the PDFs' parameters in terms of the time to peak, peak discharge, and the duration of the direct runoff hydrograph, also called the time base [*Ramirez*, 2000]. However, neither *Yue et al*. [2002] nor *Nadarajah* [2007] provide a tool to move from the dimensionless shape of the design hydrograph to an actual design flood hydrograph. *Serinaldi and Grimaldi* [2011] overcame this deficiency by linking the dimensionless hydrograph in the form of a Beta density ($f(t)$) with the design variables flood volume $V_T$ and duration $D_T$ corresponding to a fixed return period $T$ to obtain a SDH called $Q_T(t)$, expressed as:

$$Q_T(t) = f(t) V_T / D_T. \tag{1}$$

In this study, we make use of the concepts outlined above but implement a joint design event which takes account of the dependence between peak discharge and flood volume to upscale the dimensionless hydrograph shape to a synthetic design hydrograph for a given return period $T$.

The SDHs do not contain any information about the hydrological processes underlying the design event. However, additional insight into a catchment's behavior could potentially be gained when looking at design hydrographs constructed for different flood types such as flash-floods, short-rain floods, long-rain floods, or rain-on-snow floods which have different probabilities of occurrence. Such a flood type specific analysis is also advantageous from a statistical point of view because we avoid mixing events caused by different processes and can better justify the assumption commonly made in flood frequency analysis that the variables (peak discharge and flood volume) are independent and identically distributed (i.i.d.) random variables [*Klemes*, 2000; *Merz et al*., 2014] even though they are usually jointly dependent [*Serinaldi and Kilsby*, 2013]. Besides containing the physical properties and the statistical information, such flood type specific design hydrographs indicate what types of floods might require special attention in flood-risk management.

Here, we go one step further than previous studies by not only constructing catchment specific SDHs, which sum up the overall flood behavior of a catchment, but also constructing flood type specific SDHs. This is done by complementing the statistical nature of the method with process-based knowledge stored in the flood event data. It has been pointed out by *Merz and Blöschl* [2008a] that expanding information beyond the flood sample is very useful for accurately estimating flood frequencies. Knowledge on the characteristics of a flood reflected by the flood type can lead to a better understanding of floods and might improve their

prediction accuracy. *Sikorska et al*. [2015] suggested that the classification of floods at the event level along with the information of their frequency may support flood-risk management because the effects of flooding on the inundated area will alter depending on the specific flood behavior. Therefore, we propose a way of deriving flood type specific SDHs for four frequently observed and potentially hazardous flood types, i.e., flash-floods, short-rain floods, long-rain floods, and rain-on-snow floods, using a dimensionless hydrograph represented by a probability density function and design variable quantiles estimated for a specified return period using a joint frequency analysis. The method takes into account that different types of floods are characterized by different dependence structures between peak discharges and flood volumes [*Gaal et al*., 2015; *Grimaldi et al*., 2016; *Szolgay et al*., 2015].

## 2. Data

The proposed method has been developed and tested using runoff, precipitation, and temperature data from a representative set of 39 mesoscale (catchment area 20–1700 km$^2$) catchments in Switzerland (see Figure 1 and Table 3 for a complete list). The selected catchments have hourly flow data series, which is crucial for characterizing the shape of the flood hydrographs, with a length of 17–53 years, with 50% of the catchments having a record length of 30–40 years. We selected catchments only with natural flow conditions neither altered through hydropower plants nor lake regulation or water transfers, to avoid hydrograph shapes modified by direct human impacts. We excluded highly glacierized catchments because unimpaired flow records are scarce for these. The characteristics of the 39 study catchments selected are summarized in Table 1. The median catchment size is 119 km$^2$, and there are only four large catchments with a size of more than 500 km$^2$. Most measurement stations lie below 750 m.a.s.l. and only four stations are located at altitudes higher than 1000 m.a.s.l. Most catchments have a mean elevation lower than 1500 m.a.s.l. meaning that mountainous catchments are scarce [*Viviroli and Weingartner*, 2004]. However, there are two partly glacierized catchments in the set with a degree of glaciation of 6.7% and 14.2%, respectively. The median of the maximal observed peak discharges in the study catchments is 122 m$^3$/s, while the median of the maximal specific observed peak discharges (discharge per unit area) is 947 l/(s km$^2$). The median of the maximal observed flood volumes is 1918 m$^3$, although, there are catchments with significantly higher maximal flood volumes. The median of the maximal specific flood volume (volume per unit area) is 4197 10$^3$ m$^3$/km$^2$.

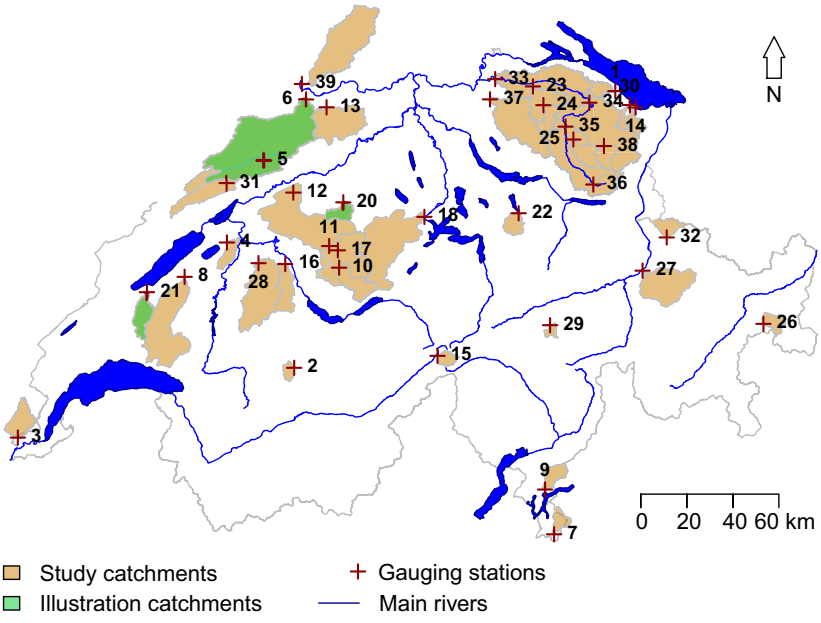



**Figure 1.** Map of the 39 catchments used for developing and testing the method for constructing flood type specific SDHs. The three catchments used to illustrate flood type specific SDHs are highlighted in green and the remaining 36 catchments in brown. The gauging stations are indicated as red crosses and labeled with the catchment ID given in Table 3.

**Table 1.** Overview of the Minimum, First Quartile, Median, Third Quartile, and Maximum of the Distribution of the Following Catchment Characteristics for the 39 Study Catchments: Record Length [years], Catchment Area [km$^2$], Station Elevation [m.a.s.l.], Mean Catchment Elevation [m.a.s.l.], Glacier Cover [%], Maximum Peak Discharge Observed in the Catchment [m$^3$/s], Maximum Specific Peak Discharge [l/(s km$^2$)], Maximum Flood Volume Observed in the Catchment [m$^3$], and Maximum Specific Flood Volume Observed in the Catchment [10$^3$ m$^3$/km$^2$]

| Catchment Characteristics | Minimum | First Quartile | Median | Third Quartile | Maximum |
|---|---|---|---|---|---|
| Record length [years] | 17 | 34 | 40 | 40 | 53 |
| Catchment area [km$^2$] | 22 | 60 | 119 | 347 | 1696 |
| Station elevation [m.a.s.l.] | 247 | 403 | 511 | 654 | 1707 |
| Mean elevation [m.a.s.l.] | 370 | 718 | 930 | 1129 | 2450 |
| Glacier cover [%] | 0 | 0 | 0 | 0 | 14 |
| Maximum peak discharge [m$^3$/s] | 11 | 62 | 122 | 276 | 956 |
| Maximum specific peak discharge [l/(s km$^2$)] | 191 | 634 | 947 | 1468 | 3123 |
| Maximum flood volume [m$^3$] | 234 | 993 | 1918 | 4503 | 27,397 |
| Maximum specific flood volume [10$^3$ m$^3$/km$^2$] | 85 | 2728 | 4197 | 6825 | 35,690 |

We selected three catchments having similar mean elevation (between 650 and 800 m.a.s.l.) to represent three catchment size classes (small, medium, and large) as examples to show more detailed results. These illustration catchments are the Langete at Huttwil (60 km$^2$), the Mentue at Yvonand (105 km$^2$), and the Birs at Münchenstein (911 km$^2$) (see green catchments in Figure 1).

## 3. Methods

The construction of flood type specific SDHs relies on the fitting of probability density functions to observed flood hydrographs of a certain flood type and takes account of the dependence between the design variables peak discharge and flood volume. The different steps involved in constructing flood type specific SDHs are displayed in Figure 2 and are described in more detail in the following paragraphs whose numbers correspond to the numbers used in Figure 2. The proposed method can either be applied to construct a catchment specific SDH without a differentiation between flood types or to construct flood type specific SDHs.

### 3.1. Flood Sampling

The SDH construction method is based only on observed runoff data. Thus, historical flood event hydrographs were selected in the 39 representative Swiss catchments described in section 2 (see Figure 1 and Table 3). To sample flood events, we used a peak-over-threshold (POT) approach based on the procedure proposed by *Lang et al*. [1999] which is superior to annual maxima sampling [*Tanaka and Takara*, 2002] because it allows for a more rational selection of events to be considered as floods than annual maxima sampling [*Lang et al*., 1999]. The threshold for the peak discharge was chosen iteratively to fulfill a target condition of four events per year on average which is a trade-off between maximizing the information content in the sample and keeping the assumption of independence between events. The independence between successive events was additionally ensured by prescribing a minimum time interval of 72 h between them. However, most of the sampled events were separated by at least 5 days. According to extreme value theory, POT values follow a generalized Pareto distribution (GPD) [*Coles*, 2001]. Therefore, we used a GPD to fit the peak discharge values. For each of these events, sampled according to the flood peaks, the flood volume was determined over a fixed event window of 72 h. The flood volumes were not selected as peak over threshold values, and did not necessarily represent annual maxima. Therefore, we tested several statistical distributions to fit the distribution of the flood volumes. The generalized extreme value distribution (GEV) was found to fit the data best. The GPD model has three continuous parameters: a location parameter $\mu$ in IR, a scale parameter $\sigma > 0$, and a shape parameter $\xi$ in IR. It is defined as:

$$F_X(x) = 1 - \left\{1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}} \quad \xi \neq 0, \tag{2}$$

where $x$ is larger than a threshold $\mu$.

On the other hand, the GEV uses the same parameters and is expressed as:

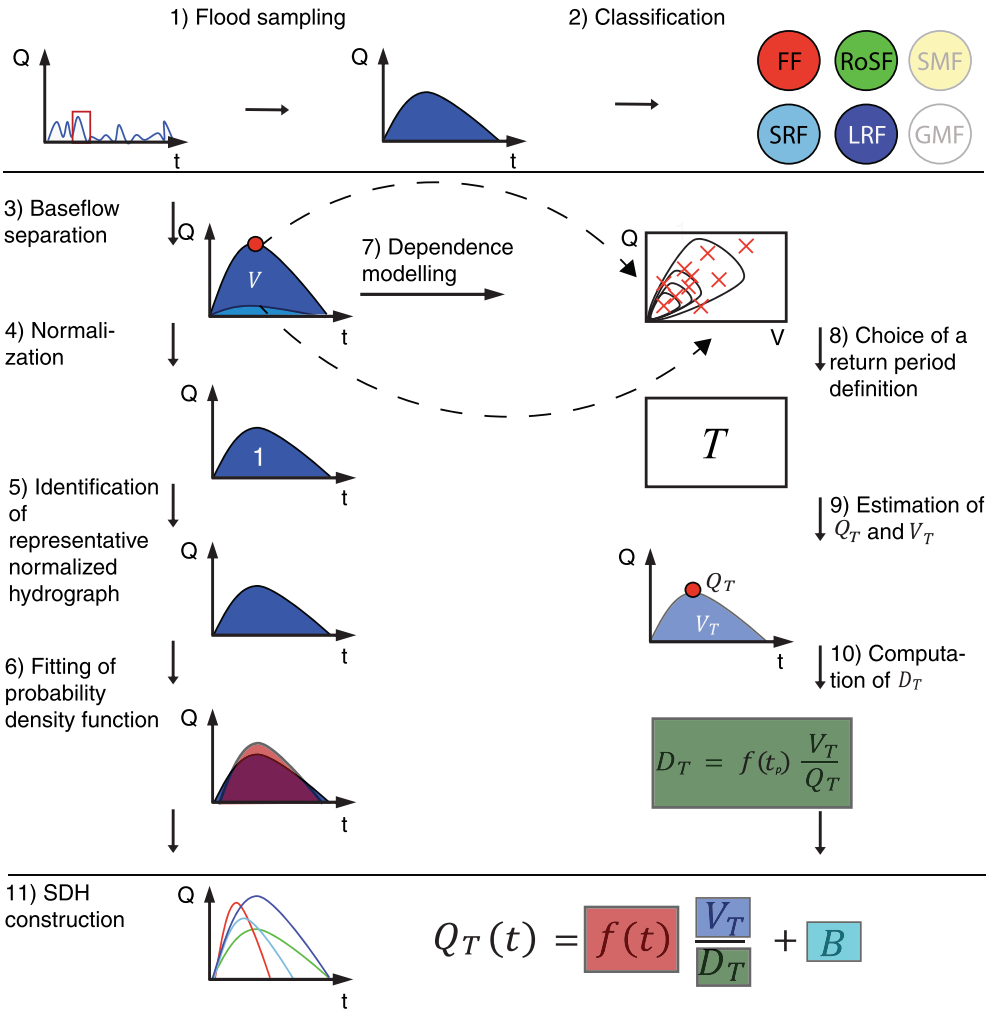



**Figure 2.** Method developed to construct synthetic design hydrographs (SDHs) for catchments in Switzerland. The method can either be applied to all types of flood events identified in a catchment or it can be used to construct flood type specific SDHs. Flood type specific SDHs can be constructed for flash-floods (FF), short-rain floods (SRF), long-rain floods (LRF), or rain-on-snow floods (RoSF), but not for snowmelt floods (SMF) and glaciermelt floods (GMF). The approach consists of 11 steps whose numbers correspond to the section numbers in the Methods chapter. The SDH ($Q_T(t)$) can be expressed by a probability density function ($f(t)$) times the mean discharge ($V_T/D_T$) plus the base flow ($B$).

$$F_Y(y) = \exp\left[-\left\{1+\xi\left(\frac{y-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}}\right] \quad \xi \neq 0 \tag{3}$$

with domain $1+\xi\left(\frac{y-\mu}{\sigma}\right) > 0$ for $\xi \neq 0$.

The goodness of fit of the GPD to the peak discharges and the GEV to the flood volumes was checked visually using *qq*-plots and *pp*-plots and tested using the Kolmogorov-Smirnov and the Anderson-Darling tests, plus the upper-tail Anderson-Darling test [*Chernobai et al.*, 2015], which gives a higher weight to the fit in the upper tail (of special interest here) than to the remaining parts of the distribution. The upper-tail Anderson-Darling test confirmed the visual impression that the GPD fits the peak discharges and the GEV distribution fits the flood volumes well.

### 3.2. Classification

To construct flood type specific SDHs, the data set needs to be divided into subsets of different flood types. Therefore, each sampled flood event was attributed to a specific flood type according to its triggering

mechanism using a flood classification scheme proposed by *Merz and Blöschl* [2003], complemented by *Diezig and Weingartner* [2007], and extended by *Sikorska et al*. [2015]. The classification scheme applied is based on successive binary splittings of the set of flood events into smaller groups of flood events according to predefined decision attributes and results in the assignment of each flood event to one of six flood types [*Sikorska et al*., 2015]. The following six flood types, having different causative mechanisms, are considered:

1. Flash-floods (FF) with short but very intense rainfall, usually lasting less than half a day.
2. Short-rain floods (SRF) with rainfall usually lasting no longer than 1 day.
3. Long-rain floods (LRF) with rainfall lasting several days or even weeks.
4. Rain-on-snow floods (RoSF) with rainfall falling on snow, which initiates its melting.
5. Snowmelt floods (SMF) caused by a melting of a snow cover with no or insignificant rainfall.
6. Glacier-melt floods (GMF) caused by glacier melting with no or insignificant rainfall.

All these flood types show a specific behavior in terms of spatial and temporal characteristics. This makes it possible to distinguish them based on a set of predefined, flood-specific indices. As proposed by *Sikorska et al*. [2015], we used the following eight indices to attribute the observed flood events to one of the six flood types: timing of the flood within the year, precipitation amount, precipitation duration, precipitation intensity, glacier cover, snow cover, snowmelt, and catchment wetness. The first four indices were computed on the basis of runoff and precipitation data as in *Sikorska et al*. [2015]. *Sikorska et al*. [2015] used the conceptual hydrological model HBV [*Seibert*, 1999] to compute glacier and snow cover, snowmelt, and catchment wetness. Here, we computed glacier cover from land cover maps, applied a simple degree-day model [*Schreider et al*., 1997] with a fixed degree-day factor (1.5 mm/degree-day) to compute the snow cover and the snowmelt, and defined the catchment wetness via the current precipitation index (CPI) [*Smakhtin and Masse*, 2000] (using a daily recession coefficient of 0.9). We tested both the crisp and the fuzzy approaches proposed by *Sikorska et al*. [2015], which led to approximately the same classification results when comparing the result from the crisp decision tree to the dominant flood type obtained with the fuzzy tree. Therefore, the decision attributes were defined as sharp thresholds to attribute exactly one flood type to each event which is facilitating computations in the construction of flood type specific SDHs. The thresholds for some indices were slightly modified from *Sikorska et al*. [2015]: The threshold for snow cover was set to a snow water equivalent of 1 mm to distinguish between the existence and nonexistence of snow cover and the threshold for the catchment wetness was set to 90% of the mean of the catchment's CPI.

Knowing all eight indices for each sampled flood event, we were able to attribute each flood event to one of the six flood types introduced above by following a decision tree with sharp thresholds. For a detailed overview of the classification scheme employed, we refer to *Sikorska et al*. [2015].

An objective validation of the flood classification process is not possible because there are no true classes [*Merz and Blöschl*, 2003]. Instead, the observed hydrographs were inspected visually and compared to the class assignments. This comparison showed that the classification procedure results in a reasonable subdivision of events into six flood types. As pointed out by *Merz and Blöschl* [2003], such an automated classification procedure is very useful in the flood frequency context because a large number of events can be classified in a limited amount of time.

The most often observed flood types in the 39 study catchments were SRFs followed by FFs, LRFs, and RoSFs (Figure 3). SMFs and GMFs were relatively rare because there are only a few mountainous catchments in the data set. While some of the flood types such as FFs, RoSFs, and SMFs show a high seasonality, other flood types such as SRFs and LRFs occur in all seasons. FFs mainly occur in summer or autumn, RoSFs mainly in winter and spring, and SMFs in winter and spring. The number of events observed per flood type varied between the different catchments. The number of FFs, LRFs, and RoSFs varied between less than 10 and around 55 events, while the number of SRFs lay between 33 and 86 events (Table 2).

In the following, we focus on the four flood types FFs, SRFs, LRFs, and RoSFs because the number of SMF and GMF events was too low to build a large enough sample to estimate SDHs. Furthermore, snowmelt floods were shown to have small flood peaks [*Merz and Blöschl*, 2003] and are therefore less relevant for flood management than the four flood types analyzed in detail.

### 3.3. Base Flow Separation

The SDH approach describes only the quick flow component of the event hydrograph originating from a precipitation event and does not consider the base flow component. Thus, it is necessary to distinguish

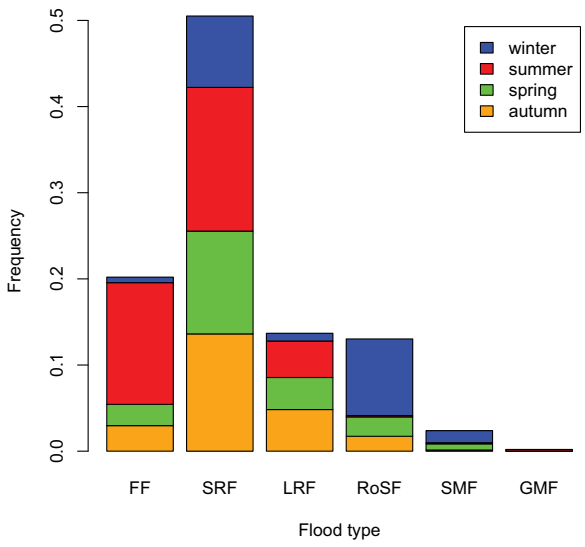



**Figure 3.** Frequency and seasonality of the six flood types obtained by the classification procedure: flash-floods (FFs), short-rain floods (SRFs), long-rain floods (LRFs), rain-on-snow floods (RoSFs), snowmelt floods (SMFs), and glacier melt floods (GMFs) over the 39 study catchments. The seasons were defined as follows: winter: December–February; spring: March–May; summer: June–August; autumn: September–November.

between the base flow and the quick runoff components to analyze the statistical properties of flood hydrographs [*Yue et al.*, 2002]. In this study, we applied a recursive digital filter [*Eckhardt*, 2005] whose two parameters need to be estimated for each catchment. This method allows for the separation of the base flow from the quick flow, is easily applicable to a wide variety of catchments, and provides reliable results in an objective way [*Serinaldi and Grimaldi*, 2011; *Gonzales et al.*, 2009]. A visual assessment showed that the method produces smooth and plausible base flow curves. In a later step, we added a representative base flow to the constructed SDH to obtain the total hydrograph. Thus, we computed a base flow index ($I_{BF}$), which is the ratio between the volume of base flow divided by the volume of total streamflow [*Smakhtin*, 2001] for each event as proposed by *Meyer et al.* [2011]. This allowed us to compute a mean base flow index for each flood type and a catchment specific base flow index.

### 3.4. Normalization

The quick flow component of the hydrographs was normalized so that both the base width and the volume of the modified hydrographs were equal to one. This was done by dividing the base width of each flood hydrograph by its duration $D$ and then dividing the ordinate of each hydrograph by the mean runoff given by the ratio of flood volume $V$ and duration $D$ ($V/D$).

### 3.5. Identification of a Representative Normalized Hydrograph

A normalized hydrograph, which is representative of a catchment's flood behavior, needs to be chosen for the nonflood-specific construction of an SDH. In addition, one representative, normalized hydrograph per flood type needs to be defined. A representative normalized hydrograph (RNH) was defined as the median normalized hydrograph of the corresponding event set (one set for all flood events and one separate set for each flood type). We chose the median normalized hydrograph instead of the mean normalized hydrograph because it refers to a real observed event, which is not necessarily the case for the mean of the normalized hydrographs.

The median hydrograph was defined using a notion of depth for functional data [*Ramsay and Silverman*, 2002]. The concept of data depth aims at measuring the centrality of a given curve (in our case, the hydrographs) within a group of curves and can be used to define the ranks of functional data [*Fraiman and Muniz*, 2001] and therewith robust estimators of a location parameter such as the median or the trimmed mean. Several data depths proposed in the literature are suitable for functional data, among them, the depth proposed by *Fraiman and Muniz* [2001], the *h*-mode depth, the random-projection depth [*Fraiman and Muniz*, 2001], or the band depth [*López-Pintado and Romo*, 2009]. We used the *h-mode* depth to order the hydrographs in the sample since it was

**Table 2.** Overview on the Minimum, First Quartile, Median, Mean, Third Quartile, and Maximum of the Number of Events Per Flood Type Over the 39 Test Catchments [the Numbers Were Rounded to Integers]

| Flood Type | Minimum | First Quartile | Median | Mean | Third Quartile | Maximum |
|---|---|---|---|---|---|---|
| FF | 4 | 15 | 19 | 24 | 34 | 59 |
| SRF | 33 | 48 | 64 | 61 | 70 | 86 |
| LRF | 5 | 10 | 13 | 16 | 20 | 45 |
| RoSF | 3 | 8 | 14 | 16 | 20 | 60 |
| SMF | 0 | 1 | 2 | 3 | 4 | 12 |
| GMF | 0 | 0 | 0 | 0 | 0 | 10 |

**Table 3.** Characteristics of 39 Swiss Catchments Used to Develop and Test the SDH Construction Approach

| ID | River | Gauging Station | Area ($km^2$) | Station Elevation (m.a.s.l.) | Mean Elevation (m.a.s.l.) | Degree of Glaciation (%) | Record Length (years) |
|---|---|---|---|---|---|---|---|
| 1 | Aach | Salmsach, Hungerbühl | 49 | 406 | 480 | 0 | 40 |
| 2 | Allenbach | Adelboden | 29 | 1297 | 1856 | 0 | 40 |
| 3 | Allondon | Dardagny, Les Granges | 119 | 400 | 758 | 0 | 29 |
| 4 | Bibere | Kerzers | 50 | 443 | 540 | 0 | 34 |
| 5 | Birse | Moutier, La Charrue | 183 | 519 | 930 | 0 | 40 |
| 6 | Birs | Münchenstein, Hofmatt | 911 | 268 | 726 | 0 | 40 |
| 7 | Breggia | Chiasso, Ponte di Polenta | 47 | 255 | 927 | 0 | 40 |
| 8 | Broye | Payerne, Caserne d'aviation | 392 | 441 | 710 | 0 | 40 |
| 9 | Cassarate | Pregassona | 74 | 291 | 990 | 0 | 40 |
| 10 | Emme | Eggiwil, Heidbüel | 124 | 745 | 1189 | 0 | 39 |
| 11 | Emme | Emmenmatt | 443 | 638 | 1070 | 0 | 17 |
| 12 | Emme | Wiler, Limpachmündung | 939 | 458 | 860 | 0 | 40 |
| 13 | Ergolz | Liestal | 261 | 305 | 590 | 0 | 40 |
| 14 | Goldach | Goldach, Bleiche | 50 | 399 | 833 | 0 | 23 |
| 15 | Goneri | Oberwald | 40 | 1385 | 2377 | 14 | 23 |
| 16 | Gürbe | Belp | 117 | 511 | 837 | 0 | 40 |
| 17 | Ilfis | Langnau | 188 | 685 | 1051 | 0 | 25 |
| 18 | Kleine Emme | Littau, Reussbühl | 477 | 431 | 1050 | 0 | 36 |
| 19 | Kleine Emme | Werthenstein, Chappelboden | 311 | 540 | 1173 | 0 | 30 |
| 20 | Langeten | Huttwil, Häberenbad | 60 | 597 | 766 | 0 | 40 |
| 21 | Mentue | Yvonand, La Mauguettaz | 105 | 449 | 679 | 0 | 40 |
| 22 | Minster | Euthal, Rüti | 59 | 894 | 1351 | 0 | 40 |
| 23 | Murg | Frauenfeld | 212 | 390 | 580 | 0 | 40 |
| 24 | Murg | Wängi | 79 | 466 | 650 | 0 | 40 |
| 25 | Necker | Mogelsberg, Aachsäge | 88 | 606 | 959 | 0 | 40 |
| 26 | Ova dal Fuorn | Zernez, Punt la Drossa | 55 | 1707 | 2331 | 0 | 40 |
| 27 | Plessur | Chur | 263 | 573 | 1850 | 0 | 40 |
| 28 | Sense | Thörishaus, Sensematt | 352 | 555 | 1068 | 0 | 36 |
| 29 | Somvixer Rhein | Somvix, Encardens | 22 | 1490 | 2450 | 7 | 36 |
| 30 | Steinach | Steinach, Mattenhof | 24 | 406 | 710 | 0 | 30 |
| 31 | Suze | Sonceboz | 150 | 642 | 1050 | 0 | 53 |
| 32 | Taschinasbach | Grüsch, Wasserfassung Lietha | 63 | 666 | 1768 | 0 | 34 |
| 33 | Thur | Andelfingen | 1696 | 356 | 770 | 0 | 40 |
| 34 | Thur | Halden | 1085 | 456 | 910 | 0 | 40 |
| 35 | Thur | Jonschwil, Mühlau | 493 | 534 | 1030 | 0 | 40 |
| 36 | Thur | Stein, Iltishag | 84 | 850 | 1448 | 0 | 31 |
| 37 | Töss | Neftenbach | 342 | 389 | 650 | 0 | 40 |
| 38 | Urnäsch | Hundwil, Äschentobel | 65 | 746 | 1085 | 0 | 33 |
| 39 | Wiese | Basel | 437 | 247 | 370 | 0 | 40 |

found to perform best if one is interested in finding the median curve within a set of curves [*Cuevas et al*., 2007].

### 3.6. Fitting of a Probability Density Function

The shape of a normalized hydrograph can be fitted by a probability density function (PDF) because both the area under the normalized hydrograph and the area under the PDF are equal to one and because probability density functions can take various shapes. *Nadarajah* [2007] and *Rai et al*. [2009] derived expressions for the unknown parameters of several density functions in terms of the time to peak ($t_p$), the peak discharge ($q_p$), and the time base ($t_B$) of the RNH. If the distribution has a finite support, its upper end point can be taken to correspond to the time base $t_B$. The time to peak $t_p$ can be defined by the value of $x$ that maximizes the PDF $f(x)$ and the peak discharge $q_p$ as the value of $f(x)$ at $t_p$. To select the most suitable density, eight different commonly used PDFs were fitted to all of the RNHs in the 39 study catchments: Normal, Lognormal, Fréchet, Weibull, Logistic, Gamma, inverse Gamma, and Beta [*Nadarajah*, 2007; *Serinaldi and Grimaldi*, 2011]. The two parameters (characterizing location and scale or scale and shape) of the distributions were estimated based on the three characteristics $t_B$, $t_p$, and $q_p$ so that the PDFs approximate the shape of the RNHs as well as possible. The goodness of fit of each PDF to the RNH was assessed by comparing the mean of the four following performance criteria for the different density functions: bounded Nash-Sutcliffe efficiency [*Mathevet et al*., 2006], volumetric efficiency, Kling-Gupta efficiency [*Gupta et al*., 2009], and Spearman's correlation coefficient. The Lognormal density function modeled the RNHs best, closely

followed by the inverse Gamma and Fréchet densities. We therefore used the Lognormal density function to model the RNHs of the different flood types and the catchment specific RNH. The modeling of the flood specific RNH was done with only a minimum of five events in the different flood type specific data sets to ensure reliable parameter estimates. This means that we could produce SDHs only for floods of those types for which we had more than five observed events in a catchment. In practice, this meant that a specific SDH for one of the flood types could not be produced in 18% of the catchments.

The fitting of PDFs to an RNH to determine the dimensionless shape of the design hydrograph proved to be effective and allows for an upscaling of the dimensionless shape to an SDH using design variable quantiles. However, the fitting of PDFs can pose a problem in the case of catchments with events with more than one peak [*Yue et al*., 2002]. There, a fitting of the hydrograph with a PDF can result in large volume differences. One could fit mixture distributions (i.e., a combination of distributions) [*Mengersen et al*., 2011] to hydrographs with multiple peaks to reduce such differences in volume but parameters of the mixture would be difficult to estimate with the sample size at hand and because of numerous interacting parameters, which would lead to identifiability problems.

### 3.7. Dependence Modeling

The dependence between peak discharge ($Q_{max}$) and flood volume ($V$) was assessed graphically using Chi-plots and K-plots and tested numerically by computing two rank correlation coefficients, Kendall's tau and Spearman's rho [*Genest and Favre*, 2007]. The bivariate distribution of peak discharges and flood volumes was expressed in the form of a copula model which, in contrast to a classical bivariate distribution, allows for modeling the dependence between the two variables independently of the choice of their marginal distributions [*Joe*, 2014]. Copulas are multivariate distribution functions whose marginal distributions are uniform. In contrast to standard multivariate distributions, copula models thus allow the variables to be characterized by different marginal distributions. The advantages of this approach are that the selection of an appropriate model for the dependence between variables, represented by the copula, can then proceed independently from the choice of the marginal distributions and that a wide selection of copula families is available to model different dependence structures [*Genest and Favre*, 2007]. For a more thorough introduction to copulas, we refer the reader to the textbooks of *Nelsen* [2005] or *Joe* [2014] or the review paper by *Genest and Favre* [2007]. The location, scale, and shape parameters of the marginal distributions GEV (equation (3)) for the flood volumes and GPD (equation (2)) for the peak discharges were estimated using the maximum likelihood method [*Coles*, 2001]. The appropriate copula to model these marginal distributions was chosen among eight copula models: five copula models of the Archimedean copula family (Gumbel, Clayton, Joe, Frank, Ali-Mikhail-Haq (AMH)), two copula models of the elliptical copula family (Student and Normal copula), and the independence copula. They were fitted to the pseudo-observations (which are deduced from the ranks of the observations) using maximum pseudolikelihood estimation, which was shown to perform best under known margins [*Hofert et al*., 2012]. After the fitting, they were tested using both graphical approaches and a goodness of fit test based on the Cramér-von Mises statistic [*Genest and Favre*, 2007]. A $p$ value for the Cramér-von Mises statistic of each copula was estimated using a statistical bootstrap procedure [*Genest et al*., 2009]. The copula models which were not rejected at the $\alpha=0.05$ significance level in most of our study catchments were the Joe and the Gumbel copula. Between these two copulas, we chose the Joe copula for further analysis because it was rejected in only 13% of the catchments. The Joe copula is described by

$$C(u,v)=1-[(1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}]^{\frac{1}{\theta}}, \tag{4}$$

where $\theta$ is the copula parameter, $u=F_X(x)$ and $v=F_Y(y)$ are uniformly distributed between 0 and 1, and their dependence is modeled by the copula $C$.

The Joe copula is very flexible and can represent the bivariate distributions of all flood types. In addition, it is able to consider tail dependence [*Heffernan et al*., 2000] which is crucial when moving toward higher return periods. The form of the dependence represented by the copula model was used for all flood types while the copula parameter $\theta$, expressing the strength of dependence, was estimated for each flood type separately. While an individual treatment of flood types can be beneficial when modeling the dependence between peak discharges and flood volumes [*Szolgay et al*., 2015; *Gaal et al*., 2015], here, testing the suitability of other copula types for the different flood types separately was not possible since the sample at hand

was too small for such an analysis [*Genest et al*., 2009]. *Grimaldi et al*. [2016] recently suggested that the actual sample size could be increased by merging the available data with other observations collected in hydrologically similar catchments. Merging the flood type specific data sets from similar catchments might allow for the estimation of the form of the copula for the individual flood types. Still, the Joe copula seems to be able to model the dependence between peak discharges and volumes well for different flood types by choosing different copula parameters for the different dependence intensities.

### 3.8. Choice of a Return Period Definition

Before estimating the design variables, a value for the return period or recurrence interval $T$ needs to be chosen (e.g., 20, 50, or 100 years). Since we deal with not only one but two nonindependent variables, peak discharge and flood volume, we also need to choose a definition for the return period in addition to the value for $T$. In a multivariate framework, a specific return period definition needs to be chosen depending on the problem at hand [*Serinaldi*, 2015a]. Several ways of defining a multivariate return period have been proposed in the literature, which all rely on different probability concepts. Definitions use the conditional probability, the joint probability (OR or AND), or can be based on Kendall's distribution or survival function. For a comprehensive overview on the topic please refer to *Brunner et al*. [2016] and the references cited therein. Assuming that the dependence between peak discharges and flood volumes is important for a potential application of this method, we used the joint OR return period definition which takes into account the dependence between $Q_{max}$ and $V$ by relying on the probability of either $Q_{max}$ or $V$ exceeding given thresholds. The joint OR return period is defined as:

$$T(x,y)=\frac{\mu}{\Pr[X>x \vee Y>y]}=\frac{\mu}{1-F_X(x)-F_Y(y)+F_{XY}(x,y)}=\frac{\mu}{1-C(u,v)}, \tag{5}$$

where $X$ and $Y$ are random variables, $C$ is a copula, $x$ and $y$ are given thresholds, $\mu$ is the interarrival time between two successive events $u=F_X(x)$ and $v=F_Y(y)$, and $F_X$, $F_Y$, and $F_{XY}$ are the marginal and joint distribution functions of the random variables, respectively.

We based our analysis on a joint return period definition but other return period definitions could also be used. As was pointed out in *Serinaldi* [2015a] and *Brunner et al*. [2016], the choice of the return period definition to be used to estimate design variable quantiles should be chosen according to the problem at hand in practice.

### 3.9. Estimation of $Q_T$ and $V_T$

The pair of design variable quantiles, $Q_{max}$ and $V$, associated with a defined joint OR return period $T$ was estimated using the marginal distributions of the variables and the Joe copula to model their dependence. For the marginal distributions, we assumed a GEV distribution for the flood volumes and a GPD for the peak discharges. We retained the pairs $(F_X(x), F_Y(y))$ that were located along the probability level $t$ corresponding to the given return period $T$ such that $1-t=1-C(u,v)$. All the pairs $(u, v)$ that are at the same probability level $t$ are eligible because they correspond to the return period $T$. The design variable pairs were then calculated by inverting their marginal distributions $F_X$ (for peak discharges) and $F_Y$ (for flood volumes)

$$Q_{\mathrm{max}T}=F_X^{-1}(u) \tag{6}$$

and

$$V_T=F_Y^{-1}(v). \tag{7}$$

There is no unique solution of the design variables associated with the joint OR return period $T$. Instead, all the possible solutions are located along the return period level, which is a curve on a bidimensional graph with $Q_{max}$ and $V$ as coordinates. We chose the design realization on this isoline that maximized the likelihood to construct the SDH [*Salvadori et al*., 2011]. A detailed description of the estimation procedure can be found in *Brunner et al*. [2016].

### 3.10. Computation of $D_T$

We restricted our analysis to the bivariate case not considering the dependence between flood volume and duration because we consider the duration of an event to be of less interest for practitioners than peak

discharge and flood volume. Therefore, the third design variable, the duration $D_T$, follows from the estimates of $Q_T$ and $V_T$ and is defined as $D_T = f(t_p) \cdot V_T/Q_T$, where $f(t_p)$ is the Lognormal density at the time of peak $t_p$. This means that the duration $D_T$ cannot be determined independently but results from the design variable quantiles $Q_T$ and $V_T$ [*Serinaldi and Grimaldi*, 2011]. Considering the dependence between event duration and flood volume would move us to a trivariate setting, where inference is computationally more challenging [*Hofert et al.*, 2012], which requires considerably more data for reliable estimation than the bivariate case [*Klein et al.*, 2010], and which is limited by the range of the dependence structures the copula chosen can handle [*Hao and Singh*, 2016].

### 3.11. SDH Construction

An SDH can be constructed for each flood type and over all flood types using the Lognormal distribution fitted to the respective RNH and the estimates for $V_T$ and $D_T$ according to equation (1). The base flow, which was removed from total flow in Step 3 of the procedure, has to be readded to the SDH to receive total flow instead of quick flow. The base flow to be added is determined by the flood type specific base flow index computed in Step 3 of the procedure. This index is multiplied with the runoff at each time step to obtain a base flow proportional to the quick flow. Hence, the construction of the final SDH requires knowledge of ten parameters: base flow index, location and scale parameter of the Lognormal distribution, a location, scale, and shape parameter for the two marginal distributions of the peak discharges and the flood volumes as well as the parameter $\theta$ of the Joe copula.

## 4. Results

### 4.1. SDH Parameters Per Flood Type

The parameters of the SDHs were estimated for the 39 study catchments in two ways: based on the whole sample of events and using only the flood type specific event sets. The results show that different flood types are characterized by different SDH parameters (Figure 4). The parameters for the base flow index,

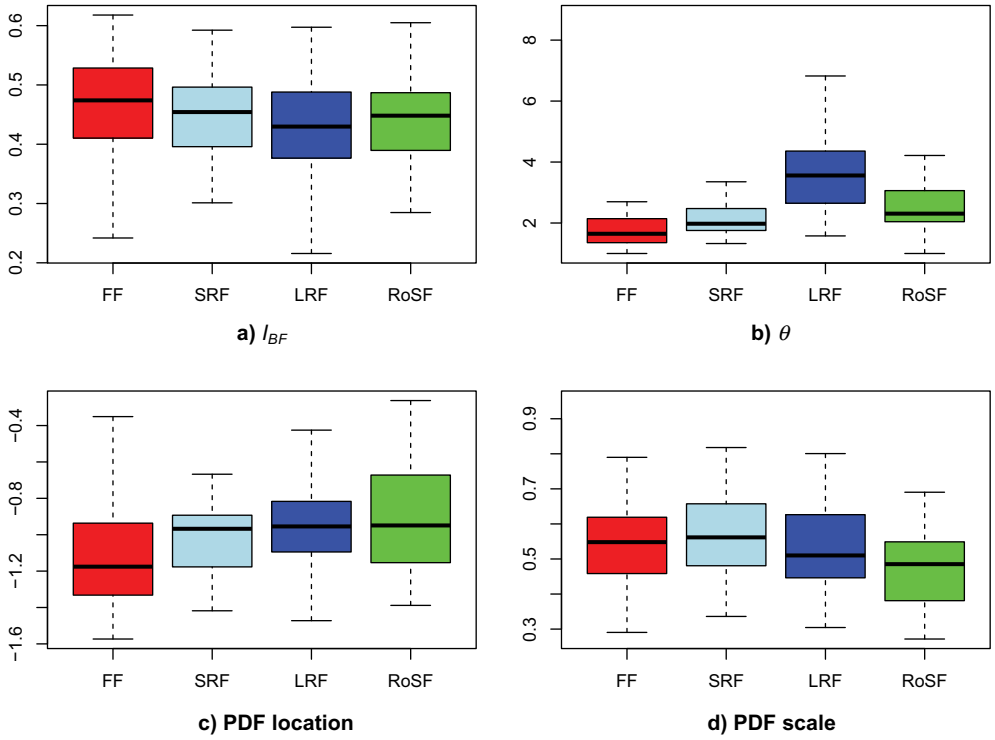



**Figure 4.** Four selected SDH parameters per flood type. (a) Base flow index ($I_{BF}$), (b) copula parameter $\theta$, (c) location parameter of the probability distribution function Lognormal (PDF location), and (d) scale parameter of the probability distribution function Lognormal (PDF scale).

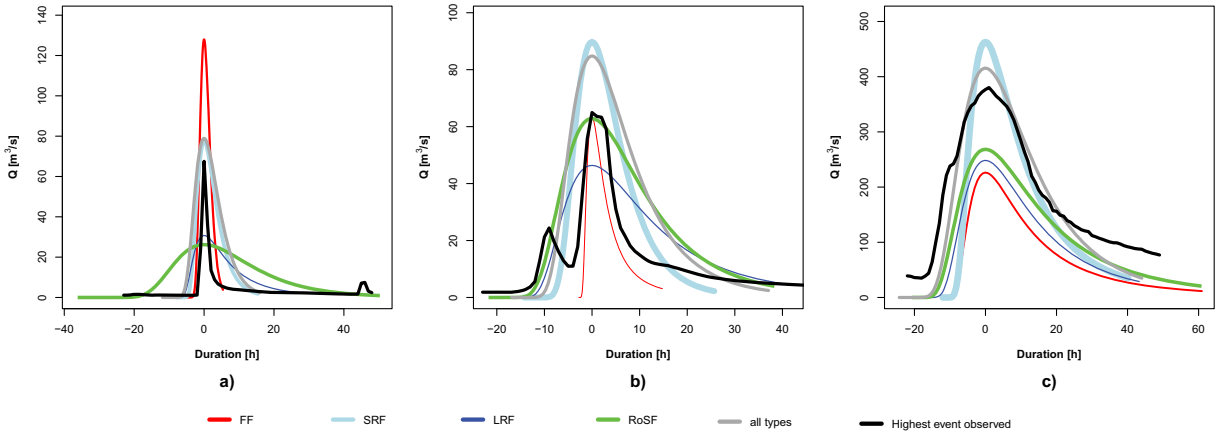



**Figure 5.** Flood type specific SDHs for three catchments of the same mean elevation zone (650–800 m.a.s.l.) but different sizes (a) Langete at Huttwil: 60 km$^2$; (b) Mentue at Yvonand: 105 km$^2$; (c) Birs at Münchenstein: 911 km$^2$. The duration is centered around the time of occurrence of the peak which was set to zero and therefore the time is negative before and positive after the peak. The line width of the SDH represents the frequency of occurrence of a certain type in the respective catchment. The highest observed event in the catchment is shown in black.

dependence ($\theta$), and the location and scale of the Lognormal PDF show some dependence on the flood type while the parameters of the marginal distributions of the peak discharges and flood volumes do not show any dependence on the flood type. The base flow index $I_{BF}$ (Figure 4a) is generally highest for FFs, followed by SRFs and RoSFs. LRFs possess the lowest base flow indices. Even though the medians of the flood type specific base flow indices differ, their variability is quite large. Concerning the copula parameter, FFs and SRFs are characterized by a low dependence between peak discharges and volumes and show thus small values of $\theta$ (Figure 4b). On the contrary, LRFs are characterized by a larger dependence between peak discharges and volumes and have higher $\theta$ values. The $\theta$ values of the RoSFs are in between those of the shorter and longer events. The location parameter of the Lognormal distribution increases with the event duration from FFs over SRFs to LRFs (Figure 4c). The scale parameter is highest for SRFs and lowest for RoSFs (Figure 4d).

### 4.2. Flood Type Specific SDHs

Based on the SDH parameters estimated per flood type, flood type specific SDHs were derived. Figure 5 displays the flood type specific SDHs for the three example catchments Langete at Huttwil (a), Mentue at Yvonand (b), and Birs at Münchenstein (c) together with an overall SDH which is not flood type specific. In the smallest catchment Langete at Huttwil, the magnitude of the peak discharges of the SDHs decreases with the duration of the precipitation event causing the flood, meaning that FFs have high peak discharges while LRFs have low peak discharges. In the medium sized catchment Mentue at Yvonand, the SRFs cause the

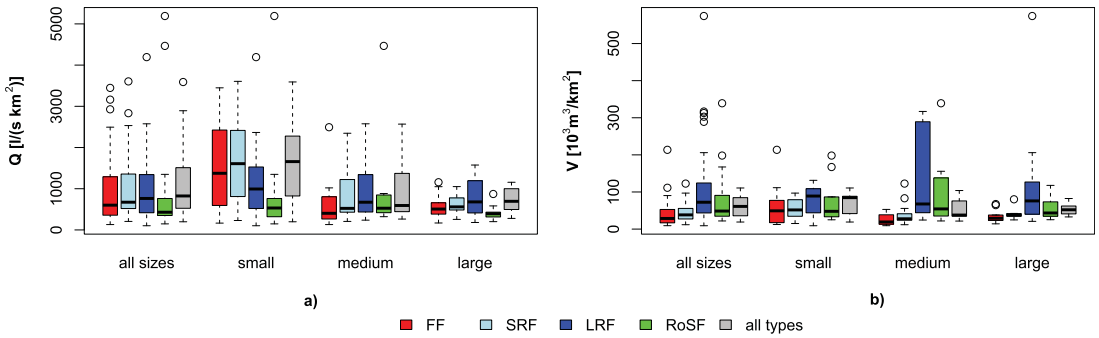



**Figure 6.** Flood type specific design variables per catchment size group: (a) specific peak discharges [l/(s km$^2$)], (b) specific flood volumes [10$^3$ m$^3$/km$^2$]. The size of small catchments ranges from 20 to 75 km$^2$, that of medium catchments from 76 to 300 km$^2$, and that of large catchments from 301 to 1700 km$^2$. The different flood types are represented by different colors.

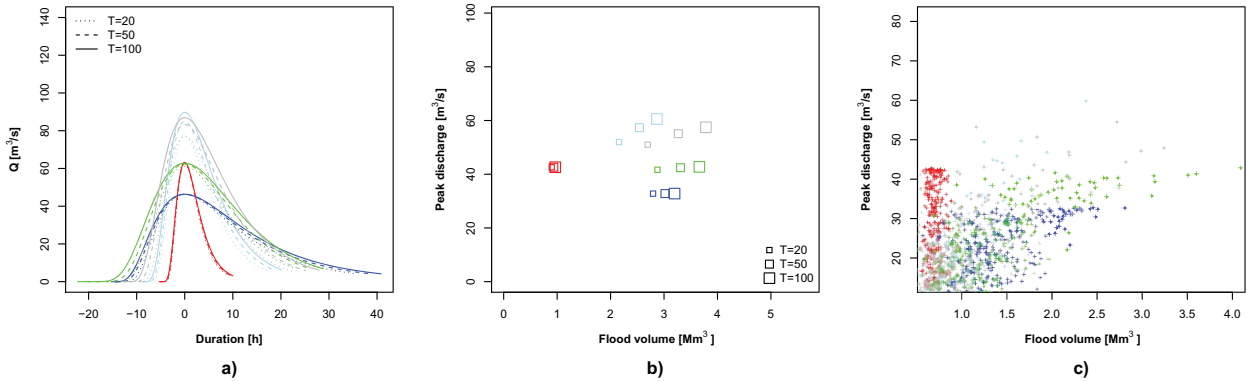



**Figure 7.** SDHs for Mentue at Yvonand for three different return periods $T = 20$, 50, 100 years for return periods defined after the joint OR probability approach for the different flood types and over all types. The joint OR probability refers to the probability that either the peak discharge, or the volume, or both, exceed a certain value. (a) Flood type specific SDHs for different return periods. (b) Flood type specific design variable quantiles for different return periods. (c) Random sample of the bivariate distribution of peak discharges and flood volumes for the different flood types and over all flood types.

highest peaks and are, together with the RoSFs, responsible for the highest volumes. In the largest catchment Birs at Münchenstein, SRFs, which are the most frequently observed event type, are characterized by high peak discharges and volumes. There is a tendency of floods being more attenuated in larger catchments than in small catchments.

Figure 6 shows the two design variables peak discharge [l/(s km$^2$)] (a) and volume [$10^3$ m$^3$/km$^2$] (b) per unit area for three different catchment size groups, namely, small catchments (20–75 km$^2$), medium sized catchments (76–300 km$^2$), and large catchments (301–1700 km$^2$), for the different flood types. When looking at catchments of all sizes, there are only small differences between specific peak discharges (peak discharge per unit area) for different flood types. However, there is a visible difference between the peaks of different flood types within the groups of catchments of similar size. In small catchments, shorter events such as FFs and SRFs have generally higher specific peak discharge than longer events such as LRFs and RoSFs. The contrary can be observed in medium sized and large catchments, where LRFs and RoSFs are generally characterized by higher specific peak discharges than shorter events such as FFs and SRFs. For the specific flood volumes, we identify a similar pattern independently of the catchment size: LRFs show higher volumes than RoSFs, and clearly higher volumes than SRFs and FFs.

### 4.3. Flood Type Specific SDHs for Different Return Periods

Figure 7 shows flood type specific SDHs (a) and flood type specific design variable quantiles (b) for three different return periods, commonly used in engineering practice, $T = 20$, $T = 50$, and $T = 100$ for the catchment Mentue at Yvonand. While the SDHs for different return levels have different peak discharges and flood volumes for most of the flood types (SRFs, LRFs, RoSFs, all types combined), the SDHs do differ only slightly for the FFs. The behavior of the different flood types for different return periods can be explained by their bivariate distribution of peak discharges and flood volumes (Figure 7c). The distribution of SRFs and of all types combined allows for both high peak discharge values and high flood volume values. The distributions of LRFs and RoSFs allow for high volume values but are bounded for the peak discharge values. The distribution of FFs is bounded in both directions. The link between different shapes of the bivariate distribution and the differences in SDHs for different return levels is confirmed in the other study catchments. However, the intensity of the dependence of the bivariate distributions of different flood types varies from catchment to catchment.

## 5. Discussion

### 5.1. SDH Parameters Per Flood Type

Different flood types were characterized by different SDH parameters. First, the base flow index was shown to depend on the total runoff volume. It can be generalized that a higher total flood volume is linked to a lower base flow index. Second, the copula parameter $\theta$, was shown to depend on the event duration. A

longer event duration is linked to a higher dependence between peak discharge and flood volume and therewith a higher $\theta$ value. Third, we showed that the shapes of shorter events, expressed by PDFs, generally show a steeper (lower scale parameter) and shorter (lower location parameter) rising limb than the shapes of longer events. The different characteristics of flood types in terms of their runoff behavior, their dependence between peak discharges and flood volumes, and their event shapes can be exploited by constructing flood type specific SDHs.

### 5.2. Flood Type Specific SDHs

The flood type specific peak discharges and flood volumes presented above show that the different flood types possess different hazard potentials. The severest floods in terms of flood volume are usually caused by LRFs independently of the size of the catchment. The flood volumes decrease from RoSFs to SRFs and FFs and are therefore linked to the event duration. The reason for this is simply that more water becomes available to form runoff with increasing duration of the precipitation event. In terms of peak discharge, on the contrary, the severest floods not only depend on the flood type but also on the catchment size. In small catchments, peak discharges decrease from shorter events such as FFs and SRFs to longer lasting events such as RoSFs and LRFs. In contrast, the peak discharges in medium sized and large catchments were higher for longer lasting events and decrease with decreasing event duration. While intense rainfall causes a fast reaction in small catchments, it is locally restricted in larger catchments [*Grebner and Roesch*, 1998] and its effect is attenuated in larger catchments on the water's way to the outlet [*Maniak*, 2010]. Convective storms are therefore more effective in small basins than in large basins [*Sutcliffe*, 1998]. In larger catchments, longer lasting events, during which the hydrological condition of the soil changes, are of importance [*Bundesamt für Wasser und Geologie* (*BWG*), 2005]. As soon as the soils are saturated, higher quantities of water will be available to form quick runoff. The hazard potential of a flood event is, besides the flood volume and the flood peak, also influenced by the shape of the design flood hydrograph [*Yue et al*., 2002; *Mediero et al*., 2010]. More storage volume is required to route a flood through a reservoir if the peak occurs early during an event than if the peak occurs later in the event. The catchment and flood type specific SDHs allow the practitioner to take account of these effects by providing information not only on flood peak and volume but also on the time of peak. The specific properties of the flood types regarding the flood hydrograph shape, design variables, and severity can help to find adequate flood protection strategies.

Knowledge on which flood type might cause the severest floods in the catchment of interest can be useful in flood prediction and flood-risk management. Some of the flood types have a pronounced seasonality and typically occur during certain seasons of the year. *Merz and Blöschl* [2003] found that in Austria, FFs mainly occur in summer and late summer while LRFs and SRFs have a less pronounced seasonality and RoSFs occur in periods when the temperature is around 0°C. Our analysis confirms this general trend for floods in Switzerland. FFs mainly occurred in summer, LRFs mainly in winter and autumn, and SRFs all year round. If a flood-risk manager knows that the catchment of interest is especially susceptible to FFs, he/she knows that the focus needs to lie on the prediction of floods caused by intense thunderstorms in summer. Flood type specific SDHs are not only helpful together with knowledge on the seasonality of occurrence but also together with the frequency of occurrence. Hydraulic structures can be laid out for the SDH of the severest and most frequent flood type in the catchment of interest. This is relevant to the practitioner because the choice of the design values directly influences the safety and the cost of a hydraulic structure [*Gräler et al*., 2013]. The protection against the severest flood to be expected in a catchment can result in oversized structures that in turn can be cost-ineffective, ecologically disadvantageous, and negative for landscape value [*Perreault and Bobée*, 1998]. Considering not only one general catchment specific SDH but four flood type specific SDHs allows the practitioner to consider the severity of a flood jointly with the frequency of occurrence of the respective flood types. This might allow for a balance between sufficient protection and feasible costs.

### 5.3. Flood Type Specific SDHs for Different Return Periods

We showed that the difference of SDHs for smaller and larger return periods depends on the bivariate distribution of peak discharges and flood volumes of the flood type analyzed. If the bivariate distribution has neither a bounded support for peaks nor volumes, a larger return period result in both a higher peak discharge and a higher flood volume than a smaller return period. If, on the contrary, the distribution's support is bounded for the peaks, a larger return period does not lead to higher peaks than a smaller return period

because higher values are not possible. The same is the case for the flood volumes if the distribution's support is bounded toward higher volumes. Whether a bivariate distribution is bounded in one or two directions depends partly on the flood type. Flood types characterized by a short duration, especially FFs, have generally little potential for the peak discharges or the flood volumes to move beyond certain values. This observation is closely linked to the classification procedure applied where FFs were defined as floods with a duration of less than six hours. The behavior of floods caused by rainfall of long duration is usually opposite of this. During an event with longer duration, the potential for higher peaks and especially higher flood volumes is given and the distribution can have a heavy upper tail. Despite the general pattern of more bounded distributions for event types with shorter duration and less bounded distributions for event types with longer distributions, the individual flood types have different bivariate distributions in different catchments. Therefore, the flood type specific SDHs of different return periods vary in the catchments analyzed. This indicates that the behavior of a flood is not only linked to the storm behavior and the precipitation input but also to the watershed and infiltration characteristics of a catchment [*Singh*, 1997].

### 5.4. Method Evaluation

The catchment specific SDH, where no differentiation is made between flood types, is usually similar to the SDH of the flood type observed most often in the catchment under consideration. If the practitioner is interested in one single hydrograph estimate, he/she might therefore work with the catchment specific SDH. If he/she, however, wants to look at the spread of possible design events, the difference between the smallest and the largest flood type specific SDH can be considered. Taking into account the spread of possible events allows one to analyze not only events of the dominant flood type in a catchment but also events of flood types observed less frequently but potentially more hazardous. While the catchment specific SDH contains information on the frequency and on the magnitude of a flood event to be expected in a catchment, the flood type specific SDHs can also serve as an indicator for underlying processes. This approach satisfies the need for hydrological reasoning in the flood frequency estimation procedure as it was postulated by *Merz and Blöschl* [2008b]. The hydrological information content is not lost completely during the statistical estimation procedure but retained in the hydrograph shapes and design variable quantiles of the flood type specific SDHs. As it was already suggested by *Klemes* [1993], we tried to shed more light on the probabilities of hydrological extremes by incorporating more information on the physical basis of the phenomena and by increasing the homogeneity of the sample by splitting it into events belonging to different flood types [*Fischer et al*., 2016]. The method can be applied to return periods of up to 100 years in gauged, medium sized catchments. The application to return periods higher than 100 years is highly discouraged because the reliability of a statistical statement is closely linked to the length of the observation period [*Deutscher Verband für Wasserwirtschaft* (*DVWK*), 1999]. The method allows for the use of different return period definitions, which need to be chosen according to the problem at hand [*Serinaldi*, 2015b; *Brunner et al*., 2016], and is not restricted to the use of the joint OR return period used here.

Due to the limited sample size, it is desirable to not only communicate estimated design variable quantiles but to complement them with uncertainty bands. It was stressed by *Serinaldi* [2009] that the design variable quantiles have to be complemented with information about their uncertainty because they are provided for events whose frequency goes beyond the range that is supported by the length of the flood records [*Reed*, 2002]. In a bivariate framework, the uncertainty related to the limited sample size and the uncertainty of the marginal distributions combine with the uncertainty of the dependence structure between the two variables [*Serinaldi*, 2015b]. In our analysis, splitting the sample of flood events into subsamples for different flood types increases the uncertainty resulting from a limited sample size.

## 6. Conclusions and Perspectives

In this study, we proposed a method that is not only suitable to construct catchment specific, but also flood type specific, synthetic design hydrographs (SDHs). The approach can not only be used when observed runoff data are available but also when the analyst is able to simulate synthetic discharge using continuous rainfall-runoff modeling. It relies on the fitting of probability density functions to observed flood hydrographs of a certain flood type taking into account the dependence between the design variables peak discharge and flood volume. The method makes use of the statistical information in the flood event data and retains some of the process-based information stored in it. It thus helps to advance from a purely statistical

methods (SM) toward a method that incorporates more hydrological information. A flood type specific construction of SDHs is meaningful because flood types differ in their runoff behavior, their dependence between peak discharges and flood volumes, and their event shapes. Even though the method has been developed and tested based on Swiss catchments, its applicability is not restricted to this geographical region but also extends to gauged catchments in other regions with similar catchments and data availability. So far, the approach is only applicable in gauged catchments with runoff and precipitation records but not in ungauged catchments. However, the estimation of design variables in ungauged catchments is of great interest [*Blöschl et al.*, 2013]. Therefore, the estimation of SDHs shall, in a next step, be regionalized to ungauged catchments where runoff data are not available. Further, the uncertainty introduced in each step of the method shall be assessed through a simulation study and the design variable estimates shall be complemented with uncertainty bands. We showed that flood type specific SDHs provide information not only on flood peak and volume but also on the time of peak and the whole event hydrograph for a certain flood type. They can be helpful in flood-risk management together with knowledge on the seasonality and frequency of occurrence of different flood types.

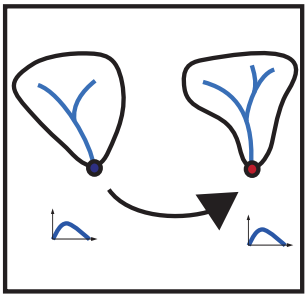
**Acknowledgments**
We thank the Federal Office for the Environment (FOEN) for funding the project (contract 13.0028.KP/M285-0623) and for providing runoff measurement data. We also thank MeteoSwiss for providing precipitation data. In addition, we thank Salvatore Grimaldi and the two anonymous reviewers for their constructive feedback, which helped to improve the quality of the paper. The data used in this study is available upon order from the FOEN and MeteoSwiss. For the hydrological data, the order form under http://www.bafu.admin.ch/wasser/13462/13494/15076/index.html?lang=de can be used. The meteorological data can be ordered via https://shop.meteoswiss.ch/index.html.

## References

Bhunya, P. K., R. Berndtsson, C. S. P. Ojha, and S. K. Mishra (2007), Suitability of Gamma, Chi-square, Weibull, and Beta distributions as synthetic unit hydrographs, *J. Hydrol.*, *334*, 28–38.
Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (2013), *Runoff Prediction in Ungauged Basins*, 465 pp., Cambridge Univ. Press, Cambridge, U. K.
Boughton, W., and O. Droop (2003), Continuous simulation for design flood estimation—A review, *Environ. Modell. Software*, *18*(4), 309–318, doi:10.1016/S1364-8152(03)00004-5.
Brunner, M. I., J. Seibert, and A.-C. Favre (2016), Bivariate return periods and their importance for flood peak and volume estimation, *Wire's Water*, *3*, 819–833, doi:10.1002/wat2.1173.
Bundesamt für Wasser und Geologie (BWG) (2005), Hydrologie der Schweiz, *Tech. Rep. 7,* Bundesamt für Wasser und Geologie, Bern.
Chernobai, A., S. T. Rachev, and F. J. Fabozzi (2015), Composite goodness-of-fit tests for left-truncated loss samples, in *Handbook of Financial Econometrics and Statistics*, edited by C.-F. Lee and J. Lee, chap. 20, pp. 575–596, Springer Sci. & Business Media, New York.
Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, 208 pp., Springer, London.
Cuevas, A., M. Febrero, and R. Fraiman (2007), Robust estimation and classification for functional data via projection-based depth notions, *Computational Statistics*, *22*(3), 481–496, doi:10.1007/s00180-007-0053-0.
Diezig, R., and R. Weingartner (2007), Hochwasserprozesstypen—Schlüssel zur Hochwasserabschätzung, *Wasser und Abfall*, *4*, 18–26.
Deutscher Verband für Wasserwirtschaft (DVWK) (1999), Statistische Analyse von Hochwasserabflüssen, *Tech. Rep*. 251, Deutscher. Verband für Wasserwirt., Bonn, Germany.
Eckhardt, K. (2005), How to construct recursive digital filters for baseflow separation, *Hydrol. Processes*, *19*, 507–515.
Fischer, S., A. Schumann, and M. Schulte (2016), Characterisation of seasonal flood types according to timescales in mixed probability distributions, *J. Hydrol.*, *539*, 38–56, doi:10.1016/j.jhydrol.2016.05.005.
Fraiman, R., and G. Muniz (2001), Trimmed means for functional data, *Test*, *10*(2), 419–440.
Gaal, L., J. Szolgay, S. Kohnova, K. Hlavcova, J. Parajka, A. Viglione, R. Merz, and G. Blöschl (2015), Dependence between flood peaks and volumes: A case study on climate and hydrological controls, *Hydrol. Sci. J.*, *60*(6), 968–984.
Genest, C., and A.-C. Favre (2007), Everything you always wanted to know about copula modeling but were afraid to ask, *J. Hydrol. Eng.*, *12*(4), 347–367.
Genest, C., B. Rémillard, and D. Beaudoin (2009), Goodness-of-fit tests for copulas: A review and a power study, *Insurance*, *44*, 199–213.
Gonzales, A. L., J. Nonner, J. Heijkers, and S. Uhlenbrook (2009), Comparison of different base flow separation methods in a lowland catchment, *Hydrol. Earth Syst. Sci.*, *13*(11), 2055–2068, doi:10.5194/hess-13-2055-2009.
Gräler, B., M. J. van den Berg, S. Vandenberghe, A. Petroselli, S. Grimaldi, B. D. Baets, and N. E. C. Verhoest (2013), Multivariate return periods in hydrology: A critical and practical review on synthetic design hydrograph estimation, *Hydrol. Earth Syst. Sci.*, *17*, 1281–1296.
Grebner, D., and T. Roesch (1998), *Flächen-Mengen-Dauer Beziehungen von Starkniederschlägen und mögliche Niederschlagsgrenzwerte in der Schweiz*, 191 pp., Geographisches Institut, ETH Zürich.
Grimaldi, S., A. Petroselli, G. Salvadori, and C. De Michele (2016), Catchment compatibility via copulas: A non-parametric study of the dependence structures of hydrological responses, *Adv. in Water Resour.*, *90*, 116–133, doi:10.1016/j.advwatres.2016.02.003.
Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, *377*, 80–91.
Hao, Z., and V. P. Singh (2016), Review of dependence modeling in hydrology and water resources, *Prog. Phys. Geogr.*, *40*(4), 549–578, doi:10.1177/0309133316632460.
Heffernan, J. E. (2000), A directory of coefficients of tail dependence, *Extremes*, *33*, 279–290.
Hofert, M., M. Mächler, and A. J. McNeil (2012), Likelihood inference for Archimedean copulas in high dimensions under known margins, *J. Multivariate Analy.*, *110*, 133–150, doi:10.1016/j.jmva.2012.02.019.
Joe, H. (2014), *Dependence Modeling With Copulas*, 480 pp., Chapman & Hall/CRC, London.
Klein, B., M. Pahlow, Y. Hundecha, and A. Schumann (2010), Probability analysis of hydrological loads for the design of flood control systems using copulas, *J. Hydrol. Eng.*, *15*(5), 360–369, doi:10.1061/(ASCE)HE.1943-5584.0000204.
Klemes, V. (1993), Probability of extreme hydrometeorological events—A different approach, *Int. Assoc. Hydrol. Sci.*, *213*, 167–176.
Klemes, V. (2000), Tall tales about tails of hydrological distributions I, *J. Hydrol. Eng.*, *5*(3), 227–231.
Lang, M., T. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling, *J. Hydrol.*, *225*, 103–117.
López-Pintado, S., and J. Romo (2009), On the concept of depth for functional data, *J. Am. Stat. Assoc.*, *104*, 718–734, doi:10.1198/jasa.2009.0108.
Maniak, U. (2010), *Hydrologie und Wasserwirtschaft*. Eine Einführung für Ingenieure, 686 pp., Springer, Berlin.

Mathevet, T., C. Michel, V. Andréassian, and C. Perrin (2006), A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins, *Int. Assoc. Hydrol. Sci.*, *307*, 211–219.
Mediero, L., A. Jiménez-Alvarez, and L. Garrote (2010), Design flood hydrographs from the relationship between flood peak and volume, *Hydrol. Earth Syst. Sci.*, *14*, 2495–2505.
Mengersen, K., C. Robert, and D. M. Titterington (2011), *Mixtures: Estimation and Applications*, 1st ed., 330 pp., John Wiley, Chichester, U. K.
Merz, B., et al. (2014), Floods and climate: Emerging perspectives for flood risk assessment and management, *Nat. Hazards Earth Syst. Sci.*, *14*(7), 1921–1942.
Merz, R., and G. Blöschl (2003), A process typology of regional floods, *Water Resour. Res.*, *39*(12), 1340, doi:10.1029/2002WR001952.
Merz, R., and G. Blöschl (2008a), Flood frequency hydrology. 1: Temporal, spatial, and causal expansion of information, *Water Resour. Res.*, *44*, W08432, doi:10.1029/2007WR006744
Merz, R., and G. Blöschl (2008b), Flood frequency hydrology. 2: Combining data evidence, *Water Resour. Res.*, *44*, W08433, doi:10.1029/2007WR006745.
Meyer, R., B. Schädler, D. Viviroli, and R. Weingartner (2011), Die Rolle des Basisabflusses bei der Modellierung von Niedrigwasserprozessen in Klimaimpaktstudien, *Hydrol. Wasserbewirt.*, *55*, 244–257.
Nadarajah, S. (2007), Probability models for unit hydrograph derivation, *J. Hydrol.*, *344*, 185–189.
Nelsen, R. B. (2005), *An Introduction to Copulas*, 2nd ed., 269 pp., Springer Sci. & Business Media, New York.
Pathiraja, S., S. Westra, and A. Sharma (2012), Why continuous simulation? The role of antecedent moisture in design flood estimation, *Water Resour. Res.*, *48*, W06534, doi:10.1029/2011WR010997.
Perreault, L., and B. Bobée (1998), Modélisation de données hydrologiques extrêmes de crues à l'aide des lois de Halphen, in *Statistical and Bayesian Methods in Hydrological Sciences*, edited by E. Parent, et al., pp. 87–106, UNESCO, Paris.
Pilgrim, D. H., and I. Cordery (1993), Flood runoff, in *Handbook of Hydrology*, edited by D. R. Maidment, chap. 9, pp. 477–546, McGraw-Hill, New York.
Rai, R. K., S. Sarkar, and V. P. Singh (2009), Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph, *Water Resour. Manage.*, *23*, 899–929.
Ramirez, J. A. (2000), Prediction and modeling of flood hydrology and hydraulics, in *Inland Flood Hazards: Human Riparian and Aquatic Communities*, edited by E. Wohl, 53 pp., Cambridge Univ. Press, Cambridge, U. K.
Ramsay, J. O., and B. W. Silverman (2002), *Applied Functional Data Analysis: Methods and Case Studies*, 190 pp., Springer, New York, doi:10.1007/b98886.
Reed, D. W. (2002), Reinforcing flood-risk estimation, *Philos. Trans. R. Soc. London*, *360*, 1373–1387.
Rickenmann, D. (1997), Sediment transport in Swiss torrents, *Earth Surf. Processes Landforms*, *22*(10), 937–951.
Rogger, M., B. Kohl, H. Pirkl, A. Viglione, J. Komma, R. Kirnbauer, R. Merz, and G. Blöschl (2012), Runoff models and flood frequency statistics for design flood estimation in Austria—Do they tell a consistent story?, *J. Hydrol.*, *456*, 30–43, doi:10.1016/j.jhydrol.2012.05.068.
Rosbjerg, D., et al. (2013), Prediction of floods in ungauged basins, in *Runoff Prediction in Ungauged Basins. A Synthesis Across Processes, Places and Scales*, edited by G. Blöschl, et al., chap. 9, pp. 189–226, Cambridge Univ. Press, Cambridge, U. K.
Salvadori, G., C. DeMichele, and F. Durante (2011), On the return period and design in a multivariate framework, *Hydrol. Earth Syst. Sci.*, *15*, 3293–3305.
Schreider, S. Y., P. H. Whetton, A. J. Jakeman, and A. B. Pittok (1997), Runoff modelling for snow-affected catchments in the Australian alpine region, eastern Victoria, *J. Hydrol.*, *200*, 1–23.
Seibert, J. (1999), Regionalisation of parameters for a conceptual rainfall-runoff model, *Agric. For. Meteorol.*, *98–99*, 279–293.
Serinaldi, F. (2009), Assessing the applicability of fractional order statistics for computing confidence intervals for extreme quantiles, *J. Hydrol.*, *376*, 528–541.
Serinaldi, F. (2015a), Dismissing return periods!, *Stochastic Environ. Res. Risk Assess.*, *29*, 1179–1189.
Serinaldi, F. (2015b), Can we tell more than we can know? The limits of bivariate drought analyses in the United States, *Stochastic Environ. Res. Risk Assess.*, *30*(6), pp. 1691–1704.
Serinaldi, F. and S. Grimaldi (2011), Synthetic design hydrographs based on distribution functions with finite support, *J. Hydrol. Eng.*, *16*, 434–446.
Serinaldi, F. and C. Kilsby (2013), The intrinsic dependence structure of peak, volume, duration, and average intensity of hyetographs and hydrographs, *Water Resour. Res.*, *49*, 3424–3442, doi:10.1002/wrcr.20221.
Sikorska, A. E., D. Viviroli, and J. Seibert (2015), Flood type classification in mountainous catchments using crisp and fuzzy decision trees, *Water Resour. Res.*, *51*, 7959–7976, doi:10.1002/2015WR017326.
Singh, V. P. (1997), Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph, *Hydrol. Processes*, *11*, 1649–1669.
Smakhtin, V. U. (2001), Low flow hydrology: A review, *J. Hydrol.*, *240*, 147–186.
Smakhtin, V. Y. and B. Masse (2000), Continuous daily hydrograph simulation using duration curves of a precipitation index, *Hydrol. Processes*, *14*, 1083–1100.
Smithers, J. C. (2012), Methods for design flood estimation in South Africa, *Water SA*, *38*, 633–646.
Sutcliffe, J. V. (1998), Flood frequency studies using regional methods, in *Statistical and Bayesian Methods in Hydrological Sciences*, edited by E. Parent, et al., pp. 341–355, UNESCO, Paris.
Szolgay, J., L. Gaal, S. Kohnova, K. Hlavcova, R. Vyleta, T. Bacigal, and G. Blöschl (2015), A process-based analysis of the suitability of copula types for peak-volume flood relationships, *Proc. Int. Assoc. Hydrol. Sci.*, *370*, 183–188, doi:10.5194/piahs-370-183-2015.
Tanaka, S., and K. Takara (2002), A study on threshold selection in POT analysis of extreme floods, *IAHS publications: The Extremes of the Extremes: Extraordinary Flood*, (*271*), 299–304.
Viglione, A., R. Merz, and G. Blöschl (2009), On the role of the runoff coefficient in the mapping of rainfall to flood return periods, *Hydrol. Earth Syst. Sci. Discuss.*, *6*(1), 627–665, doi:10.5194/hessd-6-627-2009.
Viviroli, D., and R. Weingartner (2004), The hydrological significance of mountains: From regional to global scale, *Hydrol. Earth Syst. Sci.*, *8*(6), 1016–1029.
Xiao, Y., S. Guo, P. Liu, B. Yan, and L. Chen (2009), Design flood hydrograph based on multicharacteristic synthesis index method, *J. Hydrol. Eng.*, *14*(12), 1359–1364.
Yue, S., T. Ouarda, B. Bobée, P. Legendre, and P. Bruneau (2002), Approach for describing statistical properties of flood hydrograph, *J. Hydrol. Eng.*, *7*(2), 147–153.

# Paper III



Stochastic Environmental Research and Risk Assessment
https://doi.org/10.1007/s00477-018-1523-3

ORIGINAL PAPER

# Synthetic design hydrographs for ungauged catchments: a comparison of regionalization methods

Manuela I. Brunner[1,2] · Reinhard Furrer[3,4] · Anna E. Sikorska[1,5] · Daniel Viviroli[1,6] · Jan Seibert[1,7] · Anne-Catherine Favre[2]



**Abstract**
Design flood estimates for a given return period are required in both gauged and ungauged catchments for hydraulic design and risk assessments. Contrary to classical design estimates, synthetic design hydrographs provide not only information on the peak magnitude of events but also on the corresponding hydrograph volumes together with the hydrograph shapes. In this study, we tested different regionalization approaches to transfer parameters of synthetic design hydrographs from gauged to ungauged catchments. These approaches include classical regionalization methods such as linear regression techniques, spatial methods, and methods based on the formation of homogeneous regions. In addition to these classical approaches, we tested nonlinear regression models not commonly used in hydrological regionalization studies, such as random forest, bagging, and boosting. We found that parameters related to the magnitude of the design event can be regionalized well using both linear and nonlinear regression techniques using catchment area, length of the main channel, maximum precipitation intensity, and relief energy as explanatory variables. The hydrograph shape, however, was found to be more difficult to regionalize due to its high variability within a catchment. Such variability might be better represented by looking at flood-type specific synthetic design hydrographs.

**Keywords** Regionalization · Ungauged catchments · Design hydrographs · Flood estimation · Regression trees

## 1 Introduction

Flood estimates for a given return period are required for many engineering problems, such as the construction of retention basins and weirs or drawing hazard maps (Grimaldi and Petroselli 2015; Mediero et al. 2010; Yue and Rasmussen 2002). Two types of approaches exist for the construction of design floods for a predefined return period in catchments, where runoff information is available (Smithers 2012): rainfall-runoff modeling and flood frequency analysis. Rainfall-runoff models describe how rainfall is transferred into runoff. On the one hand, excess rainfall that becomes direct runoff is modeled by considering losses through infiltration, interception, and surface storage (e.g. via runoff coefficients or infiltration equations). On the other hand, a transfer function describes how this excess rainfall is transferred into runoff (e.g. unit hydrograph (Singh et al. 2014)). These methods are based on the assumption that the return period of the rainfall input corresponds to the return period of the resulting design

✉ Manuela I. Brunner
manuela.brunner@geo.uzh.ch

[1] Department of Geography, University of Zurich, Zurich, Switzerland

[2] Université Grenoble-Alpes, CNRS, IRD, IGE, Grenoble INP, Grenoble, France

[3] Department of Mathematics, University of Zurich, Zurich, Switzerland

[4] Department of Computational Science, University of Zurich, Zurich, Switzerland

[5] Department of Hydraulic Engineering, Warsaw University of Life Sciences, SGGW, Warsaw, Poland

[6] belop gmbh, Sarnen, Switzerland

[7] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Published online: 24 February 2018

hydrograph (Chapman and Maxwell 1996), which is not always the case (Viglione et al. 2009).

Contrary to rainfall-runoff modeling, flood frequency analysis is based solely on observed runoff data (Meylan et al. 2012) and thus overcomes this assumption. However, the focus of flood frequency analysis usually lies on the hydrograph peaks (Ahn and Palmer 2016). Such a univariate analysis is not sufficient for design tasks where the storage of the flood water is of concern. In this case, information on the whole flood hydrograph is required (Pilgrim 1986). A first step towards the construction of flood hydrographs is a bivariate flood frequency analysis, which is jointly looking at peak discharges and hydrograph volumes considering their dependence (Requena et al. 2013; Shiau et al. 2006). Bivariate flood quantiles can then be combined with a dimensionless unit hydrograph (Serinaldi and Grimaldi 2011) that describes the shape of the design event. Brunner et al. (2017b) proposed a statistical flood frequency model for the construction of synthetic design hydrographs (SDHs) which combines bivariate design quantiles with a unit hydrograph represented as a probability density function (Bhunya et al. 2011). Such SDHs provide information on the hydrograph peak, the hydrograph volume, and the entire hydrograph shape. The SDH construction model is useful in gauged catchments, where runoff information is available for design hydrograph estimation. Yet, it is not directly applicable in ungauged catchments where runoff is not measured or in catchments where the record is too short to estimate reliable statistics (Ahn and Palmer 2016; Blöschl et al. 2013). Therefore, a regionalization model is needed for describing how the SDH parameters can be transferred from gauged to ungauged catchments. A range of different data types has previously been regionalized from gauged to ungauged catchments: parameters of continuous rainfall-runoff models (see reviews by He et al. (2011) and Razavi and Coulibaly (2013)), unit hydrograph parameters (Tung et al. 1997), daily flows (Farmer 2016; Kokkonen et al. 2003; Sefton and Howarth 1998), low flows (Laaha et al. 2014; Longobardi and Villani 2008; Salinas et al. 2013), flow duration curves (Boscarello et al. 2016; Cheng et al. 2012; Sauquet and Catalogne 2011), flood quantiles (GREHYS 1996; Merz and Blöschl 2004; Ouarda et al. 2001; Skoien et al. 2006), and flood event durations (Cipriani et al. 2012). However, the identification of the most suitable regionalization method is likely context dependent (Ali et al. 2012). To our knowledge, no method has so far been proposed for the regionalization of a design flood hydrograph that is solely based on runoff observations and represents both the peak and volume of the hydrograph and its shape while providing information on the event's frequency via the return period. In this study, we therefore focus on finding an appropriate method for the regionalization of design flood hydrographs to ungauged catchments. This method can subsequently be applied by engineers to derive SDHs in catchments where runoff measurements are not available. Our specific research questions were:

1. Which is the most appropriate regionalization method for transferring SDHs from gauged to ungauged catchments?
2. Do nonlinear regression techniques, such as random forest, bagging, and boosting, perform better compared to classical regionalization techniques?
3. Which catchment characteristics are most important for the prediction of SDH parameters in ungauged catchments?

To address these questions, we tested and compared 24 suitable regionalization methods that have been proposed either in the hydrological or statistical literature. These included methods establishing a relationship between catchment characteristics and design hydrograph parameters, approaches based on spatial proximity, and methods related to the building of homogeneous regions. Besides commonly used regionalization methods such as multiple linear regression and various kriging approaches, we also tested three nonlinear regression methods, i.e., bagging, random forest, and boosting, which have so far only rarely been used in hydrological regionalization studies. Regression trees have been used by Laaha and Blöschl (2006) to regionalize low flows and bagging has been used by Shu and Burn (2004) to regionalize index floods and 10-year flood quantiles. Unlike other, commonly applied methods, nonlinear regression approaches allow the consideration of hydrological processes that are nonlinear and exhibit a high degree of spatial variability (Aziz et al. 2015).

The nonlinear regression methods tested here are tree-based. They split the space of explanatory variables into a number of regions containing observations with similar response values (Strobl et al. 2009). To make a prediction for a given observation, the mean or the mode of the observations in its region can be used. The set of splitting rules used to segment the space of explanatory variables can be summarized in a tree (Hastie et al. 2008). Tree-based methods have the advantage that the model outcome is unaffected by monotone transformations of the input data and different measurement scales among explanatory variables. Furthermore, irrelevant explanatory variables are seldomly selected. The hierarchical structure of a tree ensures that the response to one input variable depends on values of inputs higher in the tree, which allows for the automatic modeling of interactions between explanatory variables (Elith et al. 2008). A tree-based method might produce good predictions on the set used to fit the model, but is likely to overfit the data, leading to a poor test

performance and high variance (James et al. 2013). Combining a large number of trees into an ensemble (Strobl et al. 2009), using bagging, random forest, or boosting methods, leads to a smoothing of the estimated response surface, results in an improvement in prediction accuracy, and reduces variance compared to applying just one single tree (James et al. 2013). Variance can be reduced by bootstrap aggregation (bagging) which takes many random samples from the population using bootstrap techniques, builds a separate prediction model using each sample, and averages the resulting predictions (Liaw and Wiener 2002). If there is a strong explanatory variable in the set, most trees in the set of trees obtained by bagging will use this explanatory variable in the top split. Random forests overcome this problem by forcing each split to consider only a subset of the explanatory variables, which corresponds to a decorrelation of the trees (Breiman 2001; James et al. 2013). While bagged trees and random forest reduce variance compared to single trees, they cannot achieve any bias reduction because each tree is based on a bootstrap sample that has the same distribution as the original data set (Elith et al. 2008). This problem can be overcome by using boosted regression trees. Boosted regression trees combine the strengths of the two algorithms regression trees and boosting (Freund and Schapire 1996; Friedman 2001, 2002). Boosting is a forward, stagewise procedure in which models (here regression trees) are fitted iteratively to the data using appropriate methods focusing on observations modelled poorly by the existing collection of trees (Hofner et al. 2009). This stagewise procedure, where successive trees depend on previous trees, distinguishes boosting from bagging.

## 2 Data

### 2.1 Study catchments

This regionalization study was performed using runoff and catchment characteristics data from 163 Swiss catchments (Fig. 1, and the Table in the "Appendix") with a wide range of catchment characteristics and flood behaviors. The selected catchments have hourly flow series of at least 20 years in duration ranging up to 53 years. The application of the hydrograph construction procedure presented in the next paragraph is suitable for catchments with records of at least 20 years, especially, if longer return periods such as $T = 100$ years are of interest (Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall 2012). The catchments' runoff is neither altered by regulated lakes upstream or inland canals nor by urbanized areas. The catchments are small to medium-size (6–1800 km$^2$), situated between 400 and 2600 m.a.s.l. (mean elevation), and either with no or only a few areas with glaciers.

### 2.2 SDH parameters

We constructed hydrographs for the study catchments using the method for a catchment specific construction of synthetic design hydrographs (SDHs) proposed by Brunner et al. (2017b). They form the data-and validation basis for the regionalization study and their parameters are provided in Table 5 in "Appendix". The method uses observed runoff data, models the hydrograph shape using a probability density function (PDF), and considers the dependence between the design variables, peak discharge ($Q_p$) and hydrograph volume ($V$) which are used for the estimation of the bivariate design quantiles $Q_T$ and $V_T$. It consists of ten steps which are shortly summarized here:

1. Flood sampling using a peak-over-threshold approach as proposed by Lang et al. (1999). The threshold for the peak discharge was chosen iteratively to fulfill a target condition of four events per year on average which is a trade-off between maximizing the information content in the sample and keeping the assumption of independence between events. The independence between successive events was additionally ensured by prescribing a minimum time interval of 72 h between them;
2. Baseflow separation using the recursive digital filter proposed by Eckhardt (2005) whose two parameters need to be estimated for each catchment;
3. Normalization of the hydrograph so that both the base width and the volume of the modified hydrographs were equal to one;
4. Identification of the median hydrograph as a representative normalized hydrograph (RNH) using the $h$-mode depth for functional data (Cuevas et al. 2007);
5. Fitting of a lognormal probability density function (PDF) (Rai et al. 2009; Yue et al. 2002) to the RNH. The shape of a normalized hydrograph can be fitted by a probability density function (PDF) because both the area under the normalized hydrograph and the area under the PDF are equal to one and because probability density functions can take various shapes;
6. Dependence modeling between peak discharges and hydrograph volumes using the Joe copula (Genest and Favre 2007; Joe 1997). The Joe copula is described by:

$$C(u,v) = 1 - \left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right]^{\frac{1}{\theta}}, \quad (1)$$

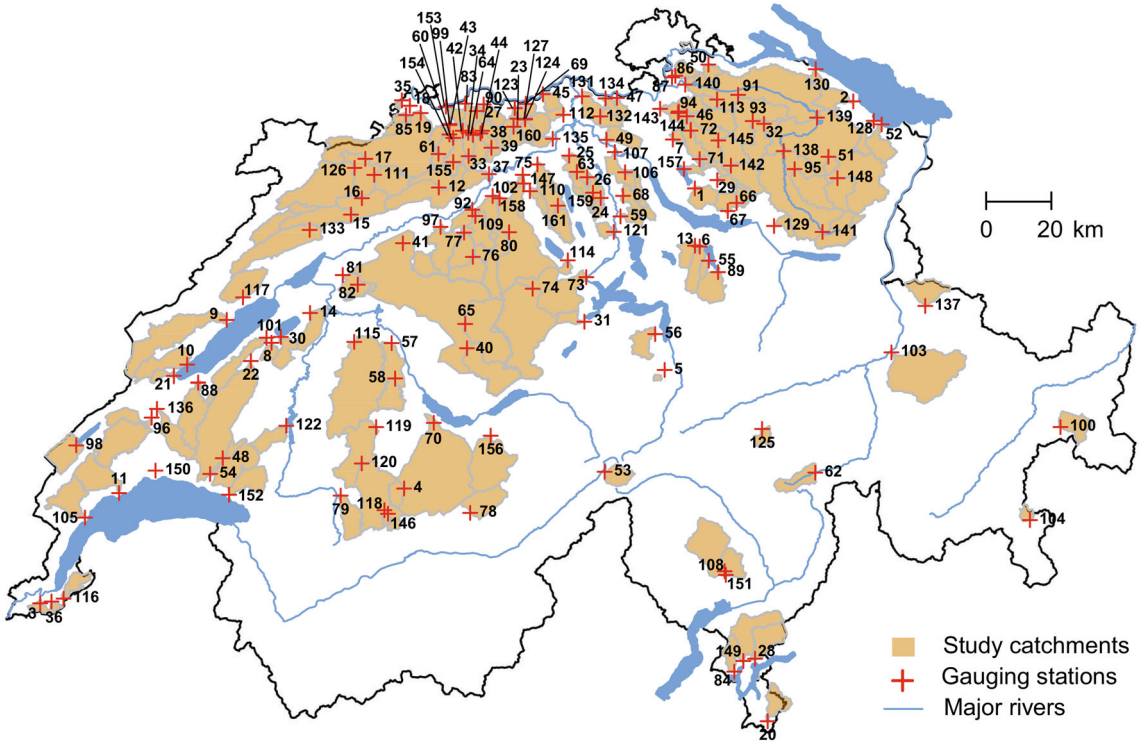



**Fig. 1** Map of the 163 Swiss study catchments used for testing different regionalization methods. The gauging stations are indicated as red crosses and labeled with the catchment ID given in the table in "Appendix"

where $\theta$ is the copula parameter, $u = F_X(x)$ and $v = F_Y(y)$ are uniformly distributed between 0 and 1, and their dependence is modeled by the copula $C$.;

7. Choice of a return period definition according to the problem at hand (Brunner et al. 2016). Here, we used the joint OR return period assuming that both peak discharge and hydrograph volume are equally important for the problem at hand. The joint OR return period is defined as follows:

$$T(x,y) = \frac{\mu_0}{Pr[X > x \vee Y > y]} = \frac{\mu_0}{1 - F_X(x) - F_Y(y) + F_{XY}(x,y)} = \frac{\mu_0}{1 - C(u,v)}, \quad (2)$$

where $X$ and $Y$ are random variables, $C$ is a copula, $x$ and $y$ are given thresholds, $\mu_0$ is the inter-arrival time between two successive events $u = F_X(x)$ and $v = F_Y(y)$, and $F_X$, $F_Y$, and $F_{XY}$ are the marginal and joint distribution functions of the random variables respectively;

8. Estimation of the design variable quantiles peak discharge ($Q_T$) and hydrograph volume ($V_T$) for the chosen return period using their marginal distributions. A marginal Generalized Pareto distribution (GPD) was used for peak discharges and a Generalized extreme value (GEV) distribution for the hydrograph volumes. The GPD model has three continuous parameters: a location parameter $\mu$ in $\mathbb{R}$, a scale parameter $\sigma > 0$, and a shape parameter $\xi$ in $\mathbb{R}$. It is defined as:

$$F_X(x) = 1 - \left\{1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}} \quad \xi \neq 0, \quad (3)$$

where $x$ is larger than a threshold $\mu$. On the other hand, the GEV uses the same parameters and is expressed as:

$$F_Y(y) = \exp\left[-\left\{1 + \xi\left(\frac{y-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}}\right] \quad \xi \neq 0 \quad (4)$$

with domain $1 + \xi\left(\frac{y-\mu}{\sigma}\right) > 0$ for $\xi \neq 0$. In the limiting case of $\xi = 0$, the GEV distribution corresponds to the Gumbel distribution. The GPD and GEV were chosen because of their good fit to the data tested with the Kolmogorov–Smirnov, Anderson–Darling, and upper-tail Anderson–Darling goodness-of-fit tests;

9. Computation of the duration of the design event ($D_T$) by $D_T = f(t_p) \cdot V_T/Q_T$, where $f(t_p)$ is the lognormal density at the time of peak $t_p$;
10. Composition of the design hydrograph using the hydrograph shape given by the PDF ($f(t)$), the design

variable quantiles ($V_T$ and $D_T$), and the baseflow ($B$) as described by:

$$Q_T(t) = f(t)V_T/D_T + B. \quad (5)$$

For a detailed description of the methodology, the reader is referred to Brunner et al. (2017b). We applied this method to construct hydrographs for a return period of 100 years, which is frequently used in hydraulic design in Switzerland (Camezind-Wildi 2005). The design flood hydrographs obtained using this method are composed of ten parameters (Fig. 2), which we herein refer to as *SDH parameters*. Two of the parameters are related to the hydrograph shape and therefore to the time evolution of the event while eight of the parameters are related to the hydrograph magnitude. The hydrograph shape is defined by the lognormal PDF with a location and a scale parameter. The marginal distributions of the design variables (peak discharges and hydrograph volumes) are modeled by three parameters each (location, scale, and shape). One parameter defines the dependence between these two variables ($\theta$) and one parameter characterizes the proportion of baseflow to be added to the direct runoff hydrograph ($I_{BF}$). The statistics of the ten SDH parameters for all study catchments are summarized in Table 1.

Some of the parameters can be assumed normally distributed (see histograms in Fig. 3), which was confirmed by the Shapiro–Wilks goodness-of-fit test (Shapiro and Wilk 1965). However, the location and scale parameters of the marginal distributions of the peak discharges (GPD) and the hydrograph volumes (GEV) are skewed to the right. The location and the scale parameter of the GPD and the GEV distribution are strictly positive (Coles 2001). Some of the SDH parameters are correlated with other SDH parameters (see scatterplots and Kendall's correlation coefficients in Fig. 3). The two parameters of the PDF are negatively correlated while the location and scale parameters of the marginal distributions GEV and GPD are positively correlated. The shape parameters of the marginal distributions of the design variables, the baseflow index ($I_{BF}$), and the dependence parameter ($\theta$) can be considered as independent of the other SDH parameters since their correlation to other SDH parameters is very weak.

The scale and shape parameters of the probability density function used to fit the runoff hydrograph are

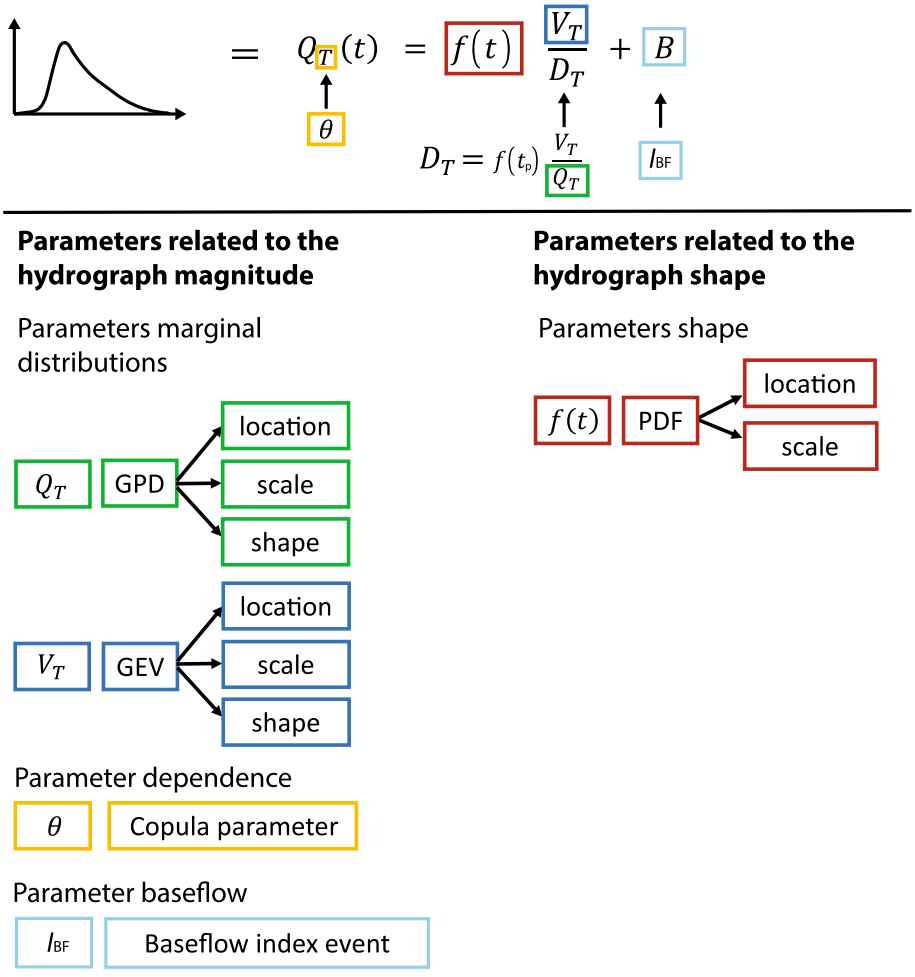



**Fig. 2** Overview of the parameters involved in the construction of an SDH. $Q_T(t)$ can be expressed as the product of the value of a PDF $f$ at the time of the peak $t_p$ and the quotient of the hydrograph volume estimated for a return period $V_T$ and the duration of the event estimated for a return period $D_T$. The baseflow ($B$) is then added to the directflow (top panel). The ten parameters needed for the construction of SDHs are presented in the bottom panel and are: Location and scale parameter of the PDF (red), parameters of the marginal distribution of the hydrograph volumes (blue), parameters of the marginal distribution of the peak discharges (green), copula parameter (yellow), and baseflow index event (light blue). The location and scale parameter of the PDF are related to the hydrograph shape while the other eight parameters are related to the hydrograph magnitude

**Table 1** Summary statistics of the ten SDH parameters over all study catchments

| SDH parameter | Mean | Standard deviation | Coefficient of variation | Unit |
|---|---|---|---|---|
| $I_{BF}$ | 0.45 | 0.13 | 0.29 | – |
| PDF location | − 0.95 | 0.21 | − 0.22 | – |
| PDF scale | 0.58 | 0.141 | 0.24 | – |
| GPD location | 13.48 | 28.56 | 2.12 | $m^3/s$ |
| GPD scale | 14.31 | 25.76 | 1.80 | $m^3/s$ |
| GPD shape | − 0.05 | 0.15 | − 3.16 | – |
| GEV location | 1.22 | 2.41 | 1.98 | $Mm^3$ |
| GEV scale | 0.42 | 0.77 | 1.84 | $Mm^3$ |
| GEV shape | 0.18 | 0.11 | 0.62 | – |
| $\theta$ | 1.88 | 0.43 | 0.23 | – |

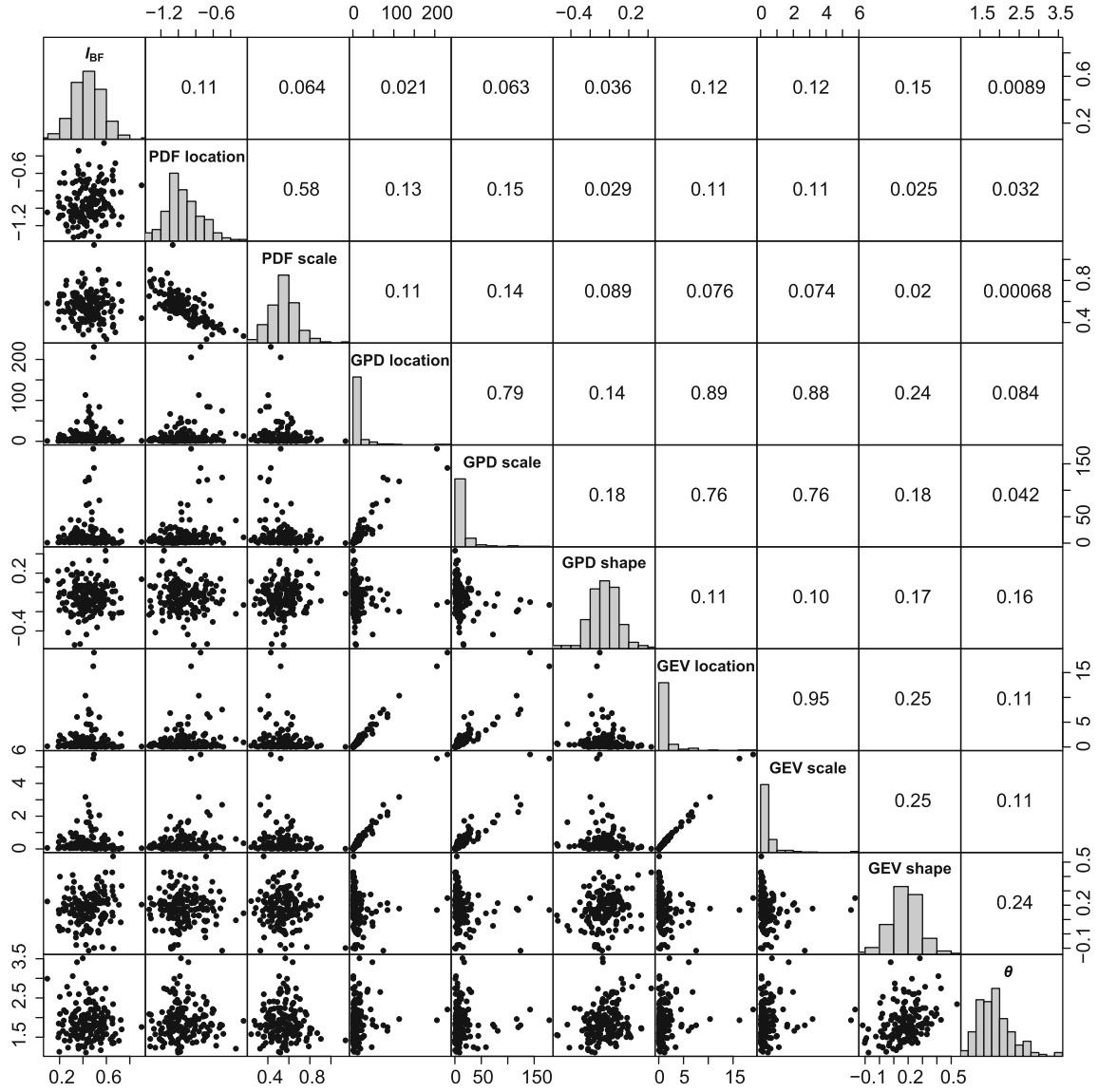



**Fig. 3** Histograms of the different SDH characteristics (diagonal), scatterplots of the SDH parameters in relation to the others (lower left panel), and Kendall's correlation coefficients between the different SDH parameters (upper right panel)

dependent. Similarly, the location, scale, and shape parameters of the GEV and the GPD are dependent on each other. Nonetheless, in the first step, these dependencies were neglected and each SDH parameter was regionalized separately. In the second step, we considered these dependencies, if allowed for by the regionalization method. Namely, multivariate regression and the methods based on the transfer of the whole parameter set from homogeneous regions, account for these dependencies.

### 2.3 Auxiliary variables

To predict runoff or runoff-related variables, such as the parameters of an SDH, in catchments without runoff measurements, alternative information has to be used (Blöschl et al. 2013) which ideally characterizes the factors that drive the hydrological response of a catchment.

*Catchment characteristics* We used a set of catchment characteristics that have been used and proven to be useful in previous regionalization studies (for a synthesis on the use of different catchment characteristics see (He et al. 2011)). Most characteristics were computed using the PREVAH pre-processing tool WINHRU (Viviroli et al. 2009b). WINHRU derived physiographical characteristics from the digital elevation model (DEM), landuse related characteristics from digital maps of landuse (Bundesamt für Statistik 2003), soil related characteristics from digital maps of land surface characteristics (Eidgenössishe Forshungsanstalt für Wald Shnee und Landshaft 1999), hydrogeology related characteristics from a map that focuses on groundwater resources (Bitterli et al. 2007), and geology related attributes from the Swiss geotechnical map (Bundesamt für Statistik 2003). For a detailed description of the computation procedure, we refer the reader to Viviroli et al. (2009a). The climatological characteristics were computed based on tables from the Hydrological Atlas of Switzerland (Jensen et al. 1997) and on gridded meteorological data provided by MeteoSwiss (MeteoSwiss 2013). The population density was computed based on a population map for Switzerland (Bundesamt für Statistik 2003).

The final set of 54 catchment characteristics (Table 2) consists of features related to the geographical location of the catchment centroid (2 characteristics), the physiography of the catchment (13), landuse (6), soil properties (6), hydrogeology (9), geology (6), climate (11), and population (1). These catchment characteristics form a solid basis for the regionalization analysis. However, some of the explanatory variables are highly correlated and contain redundant information. The potential of single catchment characteristics for aiding regionalization is indicated by an assessment of their linear relationship with the individual SDH parameters. Some SDH parameters (baseflow index ($I_{BF}$), the location and scale parameter of the marginal distributions of $Q_p$ and $V$, and the dependence parameter $\theta$) show Pearson's correlation coefficients higher than 0.4 with some of the catchment characteristics. However, the location and scale parameter of the PDF and the shape parameters of the marginal distributions of $Q_p$ and $V$ are only weakly correlated to the catchment characteristics.

*Spatial information* In addition to the characteristics mentioned above, the location of the catchments in space can be relevant for the spatial regionalization methods (see methods colored in red in Fig. 4). The coordinates of the gauging stations were used for kriging methods and a shape file for each catchment for the topological kriging methods. Linear and nonlinear approaches as well as methods based on the building of homogeneous regions use the information provided by the catchment characteristics while spatial methods use spatial information.

## 3 Methods

### 3.1 Overview

We tested several methods (Fig. 4) that have been described in hydrological and statistical literature. Testing all these methods allowed us to find the most appropriate method for the transfer of design hydrograph parameters to ungauged catchments. We grouped regionalization methods, in a similar way as other authors (e.g. Steinschneider et al. 2014), into three main classes: (1) methods based on the relation between catchment characteristics and model parameters, (2) approaches based on spatial proximity, and (3) methods based on homogeneous regions.

We followed two main strategies to regionalize the SDH parameters: (1) we regionalized each parameter individually and (2) we regionalized the ten SDH parameters (see Figs. 2, 3) as a set to assure that the relation between the parameters was not disturbed (Bardossy 2007; Parajka et al. 2005; Viviroli et al. 2009a). The individual parameters were regionalized using methods based on the relationship between catchment characteristics and model parameters, approaches based on spatial proximity, and regional mean models. On the other hand, the set of SDH parameters was regionalized using methods based on homogeneous regions. Figure 4 shows which methods were tested for the regionalization of the individual parameters and which methods were tested for the regionalization of the entire parameter set. Most models for the individual parameters were fitted on a *global* scale using data from all 163 study catchments. In addition, we fitted regional mean models for each parameter. The

**Table 2** Catchment characteristics available for regionalization and their description

| Class | Catchment characteristic | Description |
|---|---|---|
| Location | X Coordinate of catchment centroid for regression methods and of catchment outlet for kriging methods | West-East position in Swiss uniform map projection grid |
| | Y Coordinate of catchment centroid for regression methods and of catchment outlet for kriging methods | South-North position in Swiss uniform map projection grid |
| Physiography | Catchment area | – |
| | Mean altitude | – |
| | North exposed surfaces | Percentage of area exposed to the North |
| | East exposed surfaces | Percentage of area exposed to the East |
| | South exposed surfaces | Percentage of area exposed to the South |
| | West exposed surfaces | Percentage of area exposed to the West |
| | Shape parameter 1 | Area / (distance between catchment center and the most remote point)$^2$ |
| | Shape parameter 2 | Length main channel / length of main channel down to the catchment center |
| | Inclination | Percentage of surfaces with inclination $<3^\circ$ |
| | Relief energy | Maximum elevation -minimum elevation |
| | Length of main channel | – |
| | Network density | Length of main channel / area |
| | Slope of main channel | – |
| Landuse | Pastures and arable land | Percentage area with pastures and arable land |
| | Settlements | Percentage area with settlements |
| | Forest | Percentage of forested areas |
| | Glaciers | Percentage of areas with glaciers |
| | Contributing areas | Percentage or areas with a distance of 250 m to the channel |
| | Soil | Percentage of soil-covered areas |
| Soil | Hydraulic conductivity, average | Based on $100 \times 100$ m$^2$ gridded data |
| | Hydraulic conductivity, skewness | Based on $100 \times 100$ m$^2$ gridded data |
| | Net field capacity, standard deviation | Based on $100 \times 100$ m$^2$ gridded data |
| | Net field capacity, skewness | Based on $100 \times 100$ m$^2$ gridded data |
| | Soil topographic index, standard deviation | Based on $100 \times 100$ m$^2$ gridded data |
| | Soil topographic index, skewness | Based on $100 \times 100$ m$^2$ gridded data |
| Hydrogeology (groundwater) | Percentage area of unconsolidated rock, high permeability | Mainly well-porous gravel in alluvial valleys |
| | Percentage area of unconsolidated rock, intermediate permeability | Porous gravel, sandy gravel, medium-to coarse-grained debris |
| | Percentage area of unconsolidated rock, low permeability | Silty gravel, fine-to medium-grained debris, moraines |
| | Percentage area of unconsolidated rock, impermeable | Clay, silt, fine sand, loamy moraines |
| | Percentage area of hard rock, generic | Fractured-porous rock |
| | Percentage area of hard rock, impermeable | Marl, mudstone, shale, gneiss, well-cemented sandstone |
| | Percentage area of karstic rock | Carbonate and sulfate rock |
| | Hydraulic topographic index, standard deviation | Based on $100 \times 100$ m$^2$ gridded data |
| | Hydraulic topographic index, skewness | Based on $100 \times 100$m$^2$ gridded data |
| Geology (geotechnics) | Percentage area of hard rock, permeable | Various kinds of hard rock with pores, fissures, or karst |
| | Percentage area of hard rock, variable permeability | Marl, sandstone |
| | Percentage area of hard rock, impermeable | Marl, shale, clay, argillaceous slate, phyllite |
| | Percentage area of unconsolidated rock, low permeability | Clayey silts and clay |

**Table 2** Catchment characteristics available for regionalization and their description

| Class | Catchment characteristic | Description |
|---|---|---|
| | Percentage area of unconsolidated rock, variable permeability | Sands and silts, debris |
| | Percentage area of unconsolidated rock, high permeability | Gravel and sands |
| Climate | Hourly precipitation, average | Based on precipitation values $\geq$ 0.1 mm/h |
| | Hourly precipitation, coefficient of variation | Based on precipitation values $\geq$ 0.1 mm/h |
| | Hourly precipitation, seasonality | Average day of the year of occurrence |
| | Hourly precipitation, variability of seasonality | Variability of the day of the year |
| | Maximum precipitation intensity during 1 h | Return period 2.33 years |
| | Maximum precipitation intensity during 24 h | Return period 2.33 years |
| | Maximum precipitation intensity during 24 h | Average date of occurrence |
| | Maximum precipitation intensity | – |
| | Maximum precipitation intensity / Maximum precipitation intensity during 1 h | – |
| | Average annual vapor pressure | – |
| | Average annual sunshine duration | Percentage of maximum possible sunshine duration |
| Population | Population density | Number of inhabitants per area |
| Total | | 54 |

models for the regionalization of the entire parameter set were established on a regional scale and only consider a subset of these 163 study catchments. The methods tested to regionalize the ten SDH parameters are listed and shortly described below. They were implemented in *R* (R Core Team 2015) and technical details are provided in Table 3. We refer the reader to Harrell (2015) for a more detailed descriptions of linear regression methods, to Harrell (2015) or James et al. (2013) for a description of nonlinear regression methods, and to Webster and Oliver (2007) for more details on spatial methods.

## 3.2 Methods based on the relation between catchment characteristics and model parameters

A relationship was established between catchment characteristics and model parameters from the information available in gauged catchments. This relationship was assumed to be equally valid in ungauged catchments within the study region, which allowed for the estimation of model parameters from catchment characteristics. The models belonging to this class are linear and nonlinear regression models. Linear regression models only consider the linear relationship between catchment characteristics and model parameters. In contrast, nonlinear regression models such as regression trees also take into account nonlinear relationships between catchment characteristics and model parameters (Ji et al. 2013; Takezawa 2012).

### 3.2.1 Linear regression methods

We used ten multiple regression equations to express each of the SDH parameters (response variable) as a function of some catchment characteristics (explanatory variables). The regression parameters were estimated from the computed SDH parameters and the computed catchment characteristics using a least squares approach (Rosbjerg et al. 2013). As mentioned in Sect. 2, some of the SDH parameters had to be strictly positive. However, regression methods do typically not restrict prediction to strictly positive values. We followed two strategies to guarantee strictly positive values when predicting the location and scale parameters of the GPD and the GEV distribution. We worked on the log-transformed data if this guaranteed positive predictions. If not, we fitted a Gamma generalized linear model with a log-link, which does not give rise to a negative estimated response (Myers et al. 2010). For simplicity, we will refer to the linear regression models and to the Gamma generalized linear model as linear models.

*Full regression* First, we fitted a multiple regression model to the data which used all the 54 available catchment characteristics as explanatory variables. The full regression model had the disadvantage that multicollinearity was present in the set of explanatory variables, i.e., the explanatory variables were highly correlated with each other (James et al. 2013). To alleviate the problem of multicollinearity, we tested several methods that choose a subset of explanatory variables to be included in the regression model (see e.g., Harrell 2015; James et al.

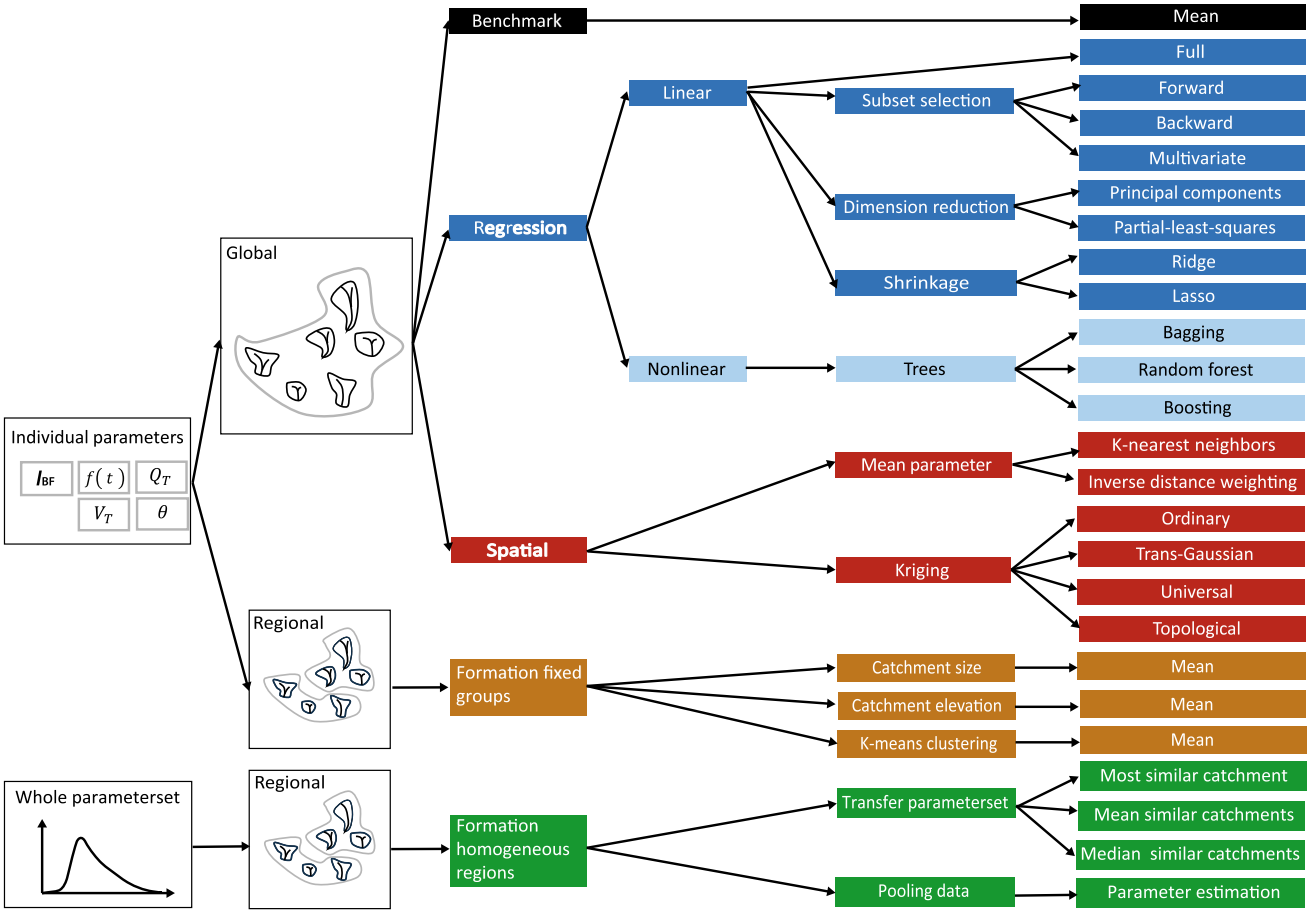



**Fig. 4** Overview of the methods tested to regionalize the ten SDH parameters on a global or regional scale. The methods were grouped into benchmark (black), regression (blue), and spatial (red) models, plus models based on the formation of regions using catchment characteristics (brown and green)

**Table 3** Technical details for the methods used in this study

| Method category | Method | R package | Functions |
|---|---|---|---|
| Linear regression | Full | *stats* (R Core Team 2015) | *lm* and *glm* |
| | Subset selection | *stats* (R Core Team 2015) | *step* |
| | Dimension reduction | *pls* (Mevik and Wehrens 2007) | *pcr* and *plsr* |
| | Shrinkage | *glmnet* (Friedman et al. 2010) | *glmnet* |
| | Multivariate | *stats* (R Core Team 2015) | *lm* |
| Nonlinear regression | Bagging | *ipred* (Peters et al. 2015) | *bagging* |
| | Random forest | *randomForest* (Liaw and Wiener 2002) | *randomForest* |
| | Boosting | *gbm* (Ridgeway 2007) | *gbm* |
| Spatial approaches | Mean parameter | *FNN* (Beygelzimer et al. 2013) | *get.knn* |
| | Kriging | *gstat* (Pebesma 2004) | *fit.variogram* and *krige* |
| | Topological kriging | *rtop* (Skoien et al. 2014) | *rtopFitVariogram* and *rtopKrige* |

2013). These can be summarized as subset selection, dimension reduction, and shrinkage techniques.

*Subset selection* Subset selection involved the identification of a subset of the 54 catchment characteristics that were most strongly related to the SDH parameters. A model was then fitted using this reduced set of variables. We applied stepwise forward and backward selection to

find a regression model with a reduced set of explanatory variables (Harrell 2015).

*Dimension reduction* We applied principal components regression (PCR) and partial-least-squares (PLS) regression (Harrell 2015; James et al. 2013; Kiers and Smilde 2007) to explain the SDH parameters by a reduced number of variables. Both methods projected the 54 explanatory variables into a smaller $m$-dimensional subspace. The regression models were then fitted using the new $m$ explanatory variables. In PCR, the regression is a linear combination of all the original explanatory variables. PLS regression is an alternative to PCR which uses not only the explanatory variables to form linear combinations but also the SDH parameters.

*Shrinkage* We used ridge regression (Le Cessie and van Houwelingen 1992) and lasso (Tibshirani 1997) to fit a model using all $p$ explanatory variables. However, the estimated regression parameters were shrunk towards zero relative to their least squares estimates. While ridge regression included all $p$ explanatory variables in the final model, the lasso shrank some of the model parameter estimates exactly to zero and therefore performed variable selection (Harrell 2015; James et al. 2013).

*Multivariate regression* The methods presented above fitted one regression equation per SDH parameter and neglected that some of the SDH parameters are correlated. Multivariate regression (or regression with multiple equations), as opposed to multiple regression, considers that some of the SDH parameters are correlated (Tung et al. 1997). We used generalized least squares (GLS) estimation, which required a covariance matrix for the errors, instead of ordinary least squares (OLS) (Weisberg 2005) to estimate the regression parameters of the multivariate regression system. The least squares residuals of the individual regression equations for each SDH parameter gave an idea of the error covariances and were used to estimate the elements of the error covariance matrix (Greene 2002). However, OLS and GLS were identical when the SDH parameters considered were uncorrelated and when all the regression equations had identical explanatory variables (Greene 2002). We therefore defined an individual subset of catchment characteristics chosen by stepwise forward selection for each SDH parameter to be used by the multivariate regression and only applied multivariate regression to those SDH parameters that were highly correlated (GEV location and scale and GPD location and scale) (see Kendall's correlation coefficients in Fig. 3). The regression parameters of the multivariate regression system were estimated using GLS. The GLS is the method to be used to fulfill all the theoretical assumptions, however, the prediction outcome is not significantly different from applying OLS on equation-by-equation regression because both OLS and GLS estimates are unbiased.

#### 3.2.2 Nonlinear regression methods

We applied three types of tree-based techniques which combine a large number of trees into an ensemble (Strobl et al. 2009): bagging, random forest, and boosting.

*Bagging* We applied bootstrap aggregation to reduce the variance and increase the prediction accuracy of regression trees (Liaw and Wiener 2002; Breiman 1996). The predictions were obtained by averaging the predictions obtained by several prediction models built based on bootstrapped samples.

*Random forest* We used random forest to decorrelate the regression trees built on several bootstrapped samples. Each split only considered a subset of the explanatory variables.

*Boosting* Boosted regression trees were used to build successive trees in a stagewise procedure where new trees depended on previous trees. Only a proportion of the observations was selected at each step to fit the tree model to prevent from overfitting.

In addition to the nonlinear approaches described above, other nonlinear methods were proposed in the literature to estimate hydrological parameters or flood quantiles for ungauged catchments. For example, Abrahart and See (2007), Aziz et al. (2015, 2016), Dawson et al. (2006), and Shu and Ouarda (2008) applied artificial neural networks (ANNs) that can be trained to represent the relationship between a range of catchment descriptors and associated hydrological parameters. However, ANNs do not allow for the determination of the role of individual variables, which reduces the confidence in model predictions (Dawson et al. 2006). They were therefore excluded from this analysis.

### 3.3 Approaches based on spatial proximity

This second class of methods is based on the assumption that the model parameters are more similar for catchments closer in a geographical or physiographical space than for catchments further apart (Rosbjerg et al. 2013). Nearly all spatial prediction methods, including the simpler forms of kriging, can be seen as weighted averages of data. Methods that are frequently used to interpolate data and use a weighted average of data from measured locations are nearest neighbor(s), inverse distance weighting, and kriging (Webster and Oliver 2007).

#### 3.3.1 Mean parameter

*Nearest neighbors* We determined the five nearest neighbors for each station in terms of the Euclidean distance and computed the mean of the different SDH parameters which then served as the predictions for the station under consideration (Parajka et al. 2005; Viviroli et al. 2009a). We

chose five neighbors as a compromise between including more information and defining non-similar catchments as nearest neighbors.

*Inverse distance weighting of nearest neighbors* We used again the five nearest neighbors of a station and gave their observations weights according to the inverse Euclidean distance from the prediction point (Hechenbichler and Schliep 2004; Lu and Wong 2008; Samuel et al. 2011). The predicted values were a weighted average of the observations at the neighboring stations.

#### 3.3.2 Kriging

In addition to the classical interpolation methods described above, we applied several kriging approaches, which consider the value observed at one point as a realisation of a spatial process (Matheron 1971). Kriging is based on the concept that values at locations near to one another are similar, whereas those at more distant locations are less correlated. This dependence structure is typically characterized by a variogram. In order to perform kriging, several steps are necessary: first, the empirical variogram is computed, second a theoretical model is fitted to this empirical variogram, then, the fitted variogram model is used to calculate the kriging weights by solving a system of equations, and finally, interpolation is carried out. The kriging methods tested differ in the assumptions about the mean structure of the model. We estimated the variograms based on the spatial information and used them for kriging. However, variograms could theoretically also be based on a physiographical space constructed using physiographical and meteorological characteristics of gauging stations and multivariate analysis techniques, such as canonical correlation analysis (CCA) or principal components analysis (PCA) (Archfield et al. 2013; Castiglioni et al. 2011; Chokmani and Ouarda 2004; Hundecha et al. 2008; Ouarda et al. 2000). In our case, many principal components or canonical variables were needed to explain an acceptable proportion of the total variance.

*Ordinary kriging* Ordinary kriging assumes that the mean of the model system is unknown but constant and that the data is symmetrically distributed. We computed the empirical variogram for all SDH parameters among which not all showed a structure (i.e., PDF location and scale showed no structure). A classical exponential variogram was fitted to the empirical variogram.

*Trans-Gaussian kriging* A symmetric distribution as assumed for ordinary kriging did not describe the behavior of all the data analyzed. One of the simplest ways to extend the symmetric model is to assume that the model holds after applying a transformation of the original data. We did trans-Gaussian kriging on the log-transformed data (Diggle and Ribeiro Jr 2007; Yamamoto 2007) for those parameters that were not normally and therefore symmetrically distributed (see Fig. 3). The kriging estimates obtained by lognormal kriging had to be back transformed to the original measurement scale by taking the exponential of the kriging estimates plus correcting the bias (Yamamoto 2007). The log-transformation was only possible for positive values (Osborne 2010).

*Universal kriging* We used universal kriging to model the spatial trend using catchment size as the explanatory variable (Bardossy and Lehmann 1997; Merz and Blöschl 2004). The variogram was computed from the residuals obtained after having removed the linear trend due to catchment size. We also tested elevation and a combination of five catchment characteristics found to be important in multiple regression as explanatory variables. However, these options led to a clearly worse performance than when using catchment area and were therefore not followed up on.

*Topological kriging* We finally used topological kriging (Skoien et al. 2006) to estimate the SDH parameters in ungauged catchments. Unlike ordinary kriging, it takes both the area and the nested nature of catchments into account (Gottschalk 1993; Gottschalk et al. 2011; Sauquet 2006).

### 3.4 Methods based on homogeneous regions

This third class of methods assumes that hydrologically homogeneous regions can be found based on catchment characteristics (Burn and Boorman 1992; Prinzio et al. 2011). We formed homogeneous regions in a physiographical space defined by catchment characteristics (Castiglioni et al. 2011; Chokmani and Ouarda 2004). In a first step, we divided the space into fixed regions (GREHYS 1996). In a second step, we formed regions of influence (Burn 1990) for each catchment separately.

#### 3.4.1 Regional mean models

We formed three regions of approximately equal size (50 catchments) that still contained enough catchments to fit a region specific regression model (Burn and Boorman 1992; Laaha et al. 2014; Ouarda et al. 2001). The regions were formed using two catchment characteristics that are considered to be hydrologically meaningful: catchment size and mean elevation. The three size zones were defined as catchments smaller than 30 $km^2$, catchments with sizes between 30 and 90 $km^2$, and catchments larger than 90 $km^2$. The three elevation zones defined were catchments lower than 600 m.a.s.l., catchments between 600 and 850 m.a.s.l., and catchments higher than 850 m.a.s.l.. These regions were used to compute the arithmetic mean (Razavi and Coulibaly 2013) and to fit regression models on the

regional scale (Nathan and McMahon 1990; Salinas et al. 2013; Sauquet and Catalogne 2011). We focused on the regional mean model since our preliminary analysis showed that regression models did not perform better on a regional than on a global scale as supported by Kjeldsen and Jones (2010) but in contrast to findings by Salinas et al. (2013) and Sauquet and Catalogne (2011). The minimum sample size to compute the arithmetic mean is much smaller than the sample size required for fitting a reliable regression model. Therefore, the number of regions could be increased. We used *k*-means clustering (Halkidi et al. 2001) to form homogeneous regions in terms of catchment characteristics (Burn 1989; Burn and Boorman 1992), which were standardized by their standard deviation over all catchments (Burn and Boorman 1993). The predictive performance of the model increased with an increase in the number of regions. However, we divided the catchments into only ten regions in order to not reduce the number of catchments per group too much, which might result in an overfitting of the model. The arithmetic mean was also computed for the regions obtained by 10-means clustering.

#### 3.4.2 Transfer from similar catchments

We first formed homogeneous regions and then transferred the parameter sets using different techniques for the transfer of the entire parameter set from similar catchments to the catchment under consideration.

*Formation of homogeneous regions* The entire set of parameters was regionalized based on the formation of regions, where each basin has its own region (Acreman and Sinclair 1986; Burn 1990) consisting of similar catchments. Similarity was defined by the Euclidean distance in the catchment characteristics space (Burn 1990; Burn and Boorman 1992; Rasmussen et al. 1993) where the catchment characteristics were normalized by their standard deviation across the whole catchment set (Oudin et al. 2010). Flood statistics were excluded from the analysis because they are not available for ungauged catchments (Ilorme and Griffis 2013). The catchment characteristics used to span the space were obtained using three different techniques since all methods for determining homogeneous regions require subjective choices (GREHYS 1996; Ilorme and Griffis 2013):

1. Hydrological reasoning (Ouarda et al. 2000): Selection of six catchment characteristics that are supposed to be hydrologically meaningful: catchment area, mean catchment elevation, network density, mean annual rainfall, and *X*-and *Y*-coordinates.
2. Best-*H*: Random set of catchment characteristics leading to the lowest mean *H*-statistic, a homogeneity measure proposed by Hosking and Wallis (1993), for the regions obtained for the different catchments. This set consists of the following catchment characteristics: proportion of south exposed areas in the catchment, length of the main channel, percentage of area covered by hard rock, percentage of surfaces with inclination smaller than three degrees, and maximum 24h-precipitation.
3. CCA: Set of catchment characteristics obtained by canonical correlation analysis (CCA) between the catchment characteristics and the *L*-moments of peak discharges and hydrograph volumes (Cavadias et al. 2001; He et al. 2011; Ouarda et al. 2000, 2001). This set consists of the following catchment characteristics: average precipitation, maximum precipitation intensity over a time interval of 1 h, percentage of contributing areas, network density, percentage of area covered by hard rock, *X*-and *Y*-coordinate, hourly precipitation variability, percentage of surfaces with inclination smaller than three degrees, population density, percentage of area covered by unconsolidated rock, percentage of soil covered areas, percentage of sealed areas, and percentage of agricultural areas.

Based on the distance matrices, the most similar catchments were identified for each target catchment. We used the five most similar catchments because this number optimized the predictive performance over all SDH parameters and has already been found appropriate in previous studies (Viviroli et al. 2009a). Still, the homogeneity in terms of Hosking and Wallis' *H*-statistic was not good. Only 26% of the regions formed could be said to be homogeneous in terms of peak discharge (looking at its *L*-moments), which is not surprising since there is often a lack of correlation between catchment descriptors and flow-derived characteristics (Ali et al. 2012).

*Transfer methods* Despite this lacking hydrological homogeneity, we transferred the SDH parameter set from the formed regions to the target catchment using the following four strategies:

1. Transfer of the parameter set from the most similar catchment adjusted by the catchment size (Parajka et al. 2005; Zhang and Chiew 2009);
2. Transfer of the mean parameter set of the five most similar catchments adjusted by the catchment size (Oudin et al. 2008). We chose the five most similar catchments since this can average out the effect of choosing a poor donor catchment and most studies use between five and ten catchments for averaging (Zhang and Chiew 2009). Averaging the parameters resulted in much more plausible results than averaging the output of the five models;
3. Transfer of the median parameter set of the five most similar catchments adjusted by the catchment size. The median set is expected to reduce the influence of a

catchment having rather different SDH parameters than the other catchments compared to the mean set;
4. Estimation of a new parameter set based on the pooled runoff data from the five most similar catchments. Data pooling increases the size of the data set (Ilorme and Griffis 2013) which can be especially advantageous when estimating the parameters of the distribution of extremes (Castellarin et al. 2001; Gaál et al. 2008).

### 3.5 Benchmark model

To compare the predictive performances of the regionalization methods tested, we used the arithmetic mean of the individual SDH parameters as a simple benchmark model (Parajka et al. 2005; Razavi and Coulibaly 2013; Steinschneider et al. 2014).

### 3.6 Model validation

We used $k$-fold cross validation to validate the performance of the regionalization methods tested. The data was divided into 10 parts, also called folds, of equal size. The model parameters were estimated based on 10 minus 1 folds, and the model was validated on the remaining fold. This procedure was repeated for each fold in turn (Hastie et al. 2008; James et al. 2013).

To validate the methods, we compared on the one hand the regionalized or predicted SDH parameters to the SDH parameters estimated based on runoff data. On the other hand, we compared the hydrograph resulting from the regionalized SDH parameters to the SDH estimated based on runoff observations. Complementing the validation of the hydrograph characteristics with a validation of the single parameters allows for a better understanding of why a given approach outperforms the other (Steinschneider et al. 2014). The performance of the different regionalization methods in terms of the individual SDH parameters can be assessed using different measures such as the bias (deviation in the mean), the mean squared error ($E_{MS}$), the root mean squared error ($E_{RMS}$), or the (absolute) relative error (Chebana and Ouarda 2009; Ouarda et al. 2006; Salinas et al. 2013; Shu and Ouarda 2008). Here, we used the absolute relative error ($E_{AR}$) and the mean absolute relative error ($E_{MAR}$) because they allow us to compare the performance of the different regionalization methods to the performance of the benchmark model. Further, they allow a comparison of how well the different SDH parameters can be predicted. $E_{AR}$ and $E_{MAR}$ (Sauquet 2006) are given by

$$E_{AR} = \left| \frac{X_{ic} - X_{ir}}{X_{ic}} \right| \tag{6}$$

and

$$E_{MAR} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{X_{ic} - X_{ir}}{X_{ic}} \right|, \tag{7}$$

where $X_{ic}$ is the computed SDH parameter or characteristic, $X_{ir}$ is the regionalized parameter or SDH characteristic, and $n$ is the number of catchments in the data set.

When looking at the predictive performance in terms of the whole hydrograph, one single goodness-of-fit criterion is usually not sufficient to assess the fit between a constructed and a regionalized hydrograph (Green and Stephenson 1986). Therefore, Green and Stephenson (1986) recommended to compute the relative error of individual hydrograph characteristics such as peak discharge ($Q_p$) and volume ($V$) between the regionalized and the constructed hydrograph. In addition, we also assessed the error of the hydrograph shape via the relative error of the time to peak ($t_p$) (Tung et al. 1997) and the time between the peak and where the recession reaches half of the peak discharge ($t_{p05}$). We will refer to $t_{p05}$ as *half-recession time* (see Fig. 5 for an illustration of the different hydrograph characteristics).

### 3.7 Importance of catchment characteristics

Several of the tested regionalization methods allow us to identify catchment characteristics that were important to explain the SDH parameters via a variance importance plot. Namely, these approaches are the subset selection techniques (forward and backward), the shrinkage technique lasso, the dimension reduction techniques principal components-and partial-least squares regression, bagged regression trees, random forest, and boosted regression trees. We identified the five most important catchment characteristics in each of these approaches. Catchment characteristics that are important in at least four out of the eight approaches were said to be important for the prediction of the SDH parameter under consideration.

## 4 Results

### 4.1 Validation of the individual SDH parameters

Figure 6 shows the $E_{MAR}$ of the ten SDH parameters for the regionalization methods that were used to regionalize the individual parameters comprising the benchmark model, linear regression models, nonlinear regression models, spatial models, and regional mean models (see Fig. 4). Our results show that we can find a better model than the benchmark model for the regionalization of most SDH parameters ($I_{BF}$; GPD location, scale, and shape; GEV location, scale, and shape; and $\theta$). However, this is not

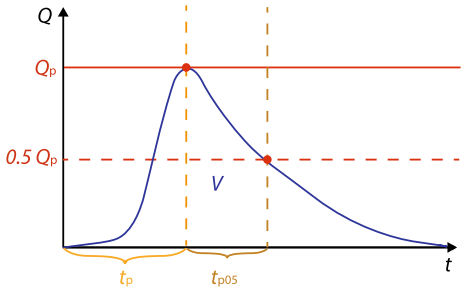



**Fig. 5** Illustration of the hydrograph characteristics used for the validation of the regionalization methods: peak discharge ($Q_p$), hydrograph volume ($V$), time to peak ($t_p$), and half-recession time ($t_{p05}$)

possible for those SDH parameters describing the hydrograph shape (PDF location and scale). This implies that working with a simple mean value for the PDF parameters is sufficient and working with a more complex model does not lead to improved results. In general, nonlinear regression techniques (especially boosting) perform slightly better than linear regression techniques. For some parameters ($I_{BF}$, GPD shape, and $\theta$), spatial methods perfom better than linear regression methods. For others (GPD location and scale and GEV location and scale) linear regression methods perform better than spatial methods (except for universal kriging where a linear component is included in the model). The regional mean models generally perform better than the global mean model (benchmark). $E_{MAR}$ varies significantly across SDH parameters. On the one hand, the parameter $I_{BF}$ and the two PDF parameters show low $E_{MAR}$ (i.e. good performance) over all regionalization methods including the benchmark model. On the other hand, the SDH parameters related to the magnitude of the event (GEV location, scale, and shape and GPD location, scale, and shape) show relatively high $E_{MAR}$ (i.e. bad performance) for some methods.

## 4.2 Validation of the whole design hydrograph

The predictive performance of the whole hydrograph (for an example see Fig. 7) is displayed in Fig. 8. It is represented by

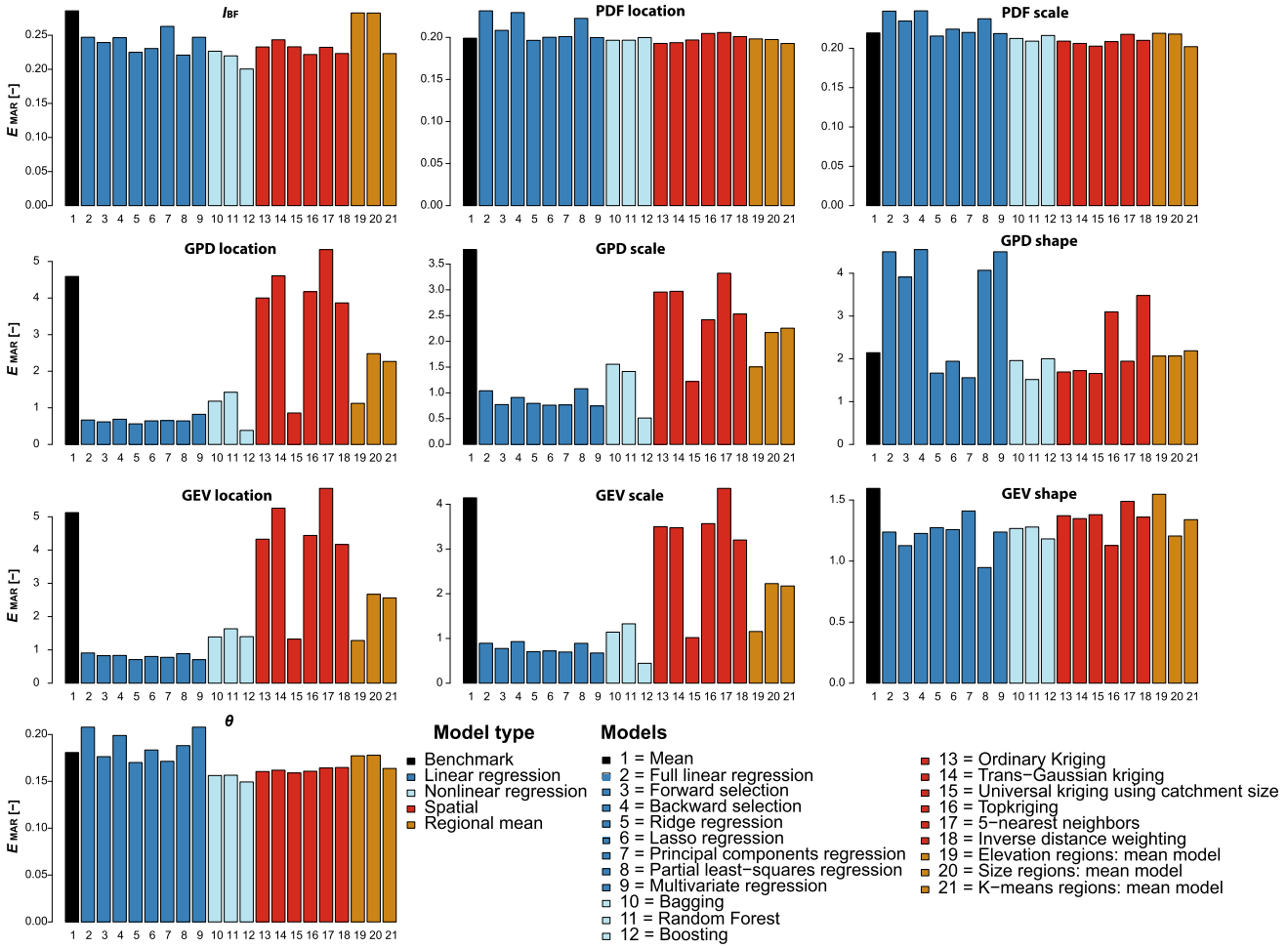



**Fig. 6** Predictive performance of the different regionalization methods (black: benchmark, blue: linear regression, light blue: nonlinear regression, red: spatial, brown: regional mean models) in terms of the mean absolute relative error ($E_{MAR}$) for the ten SDH parameters

boxplots of the $E_{AR}$ for the different catchments for the four different hydrograph characteristics $Q_p$, $V$, $t_p$, and $t_{p05}$ (see Fig. 5). The predictive performance for $Q_p$ (Fig. 8a) and $V$ (Fig. 8b) is significantly better for linear and nonlinear regression techniques and when transferring the entire parameter set based on the formation of homogeneous regions (using the strategy of hydrological reasoning) than for the benchmark model. However, when looking at the hydrograph shape represented by the time to peak and the half-recession time (Figs. 8c and d), no model can be found that better regionalizes the shape than the benchmark model. The $E_{AR}$ are generally higher for the magnitude of the event ($Q_p$ and $V$) than for the shape of the event ($t_p$, $t_{p05}$). The $E_{AR}$ of the temporal hydrograph characteristics $t_p$ and $t_{p05}$ are not correlated to the $E_{AR}$ of the magnitude of the event characterized by $Q_p$ and $V$. This means that a model badly predicting the shape of the event does not necessarily badly predict the magnitude of the event and *vice versa*. However, the performance of the two characteristics $Q_p$ and $V$ is closely linked (Kendall's correlation coefficient $> 0.6$).

#### 4.2.1 Regionalization based on the formation of homogeneous regions

The predictive performance of the regionalization methods based on the formation of homogeneous regions in terms of the four hydrograph characteristics $Q_p$, $V$, $t_p$, and $t_{p05}$ can not be improved by applying a different distance metric than the one based on hydrological reasoning (Fig. 9). The $E_{AR}$ was neither improved by the formation of homogeneous regions using catchment characteristics identified by a CCA analysis nor using catchment characteristics found to be related to hydrological characteristics via a sampling experiment (best-$H$).

### 4.3 Importance of catchment characteristics

The catchment characteristics which were important for the prediction of the ten SDH parameters in ungauged catchments are listed in Table 4. Catchment characteristics related to geology and hydrogeology were important for the prediction of the $I_{BF}$. Geology or more specifically the presence or non-presence of karstic rock was also important for the prediction of the SDH parameters related to the shape of the hydrograph (PDF location and PDF scale). On the contrary, catchment area was important for the prediction of the SDH parameters related to the magnitude of the event (GPD location, GPD scale, GEV location, and GEV scale). Exposition was meaningful for the prediction of the shape parameters of the marginal distributions of $Q_p$ and $V$ and the prediction of the dependence parameter $\theta$.

## 5 Discussion

### 5.1 Comparison of regionalization methods

The results show that it is possible to regionalize the synthetic design hydrograph (SDH) parameters representing

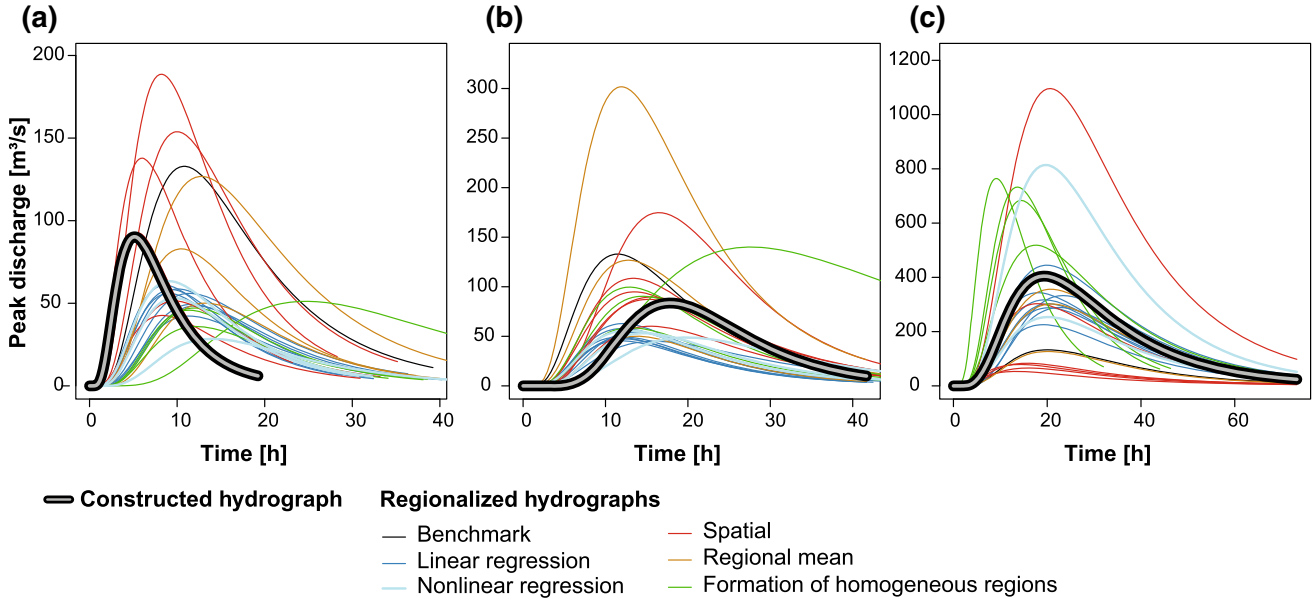



**Fig. 7** Regionalized SDHs for three example catchments of different catchment size: Langete-Huttwil (60 km$^2$), Mentue-Yvonand (105 km$^2$), and Birs-Münchenstein (911 km$^2$) for different regionalization techniques: Benchmark model (black), Linear regression models (blue), nonlinear regression models (light blue), spatial methods (red), regional mean models (brown), formation of homogeneous regions (green). The hydrograph constructed based on observed runoff data is displayed in grey on black

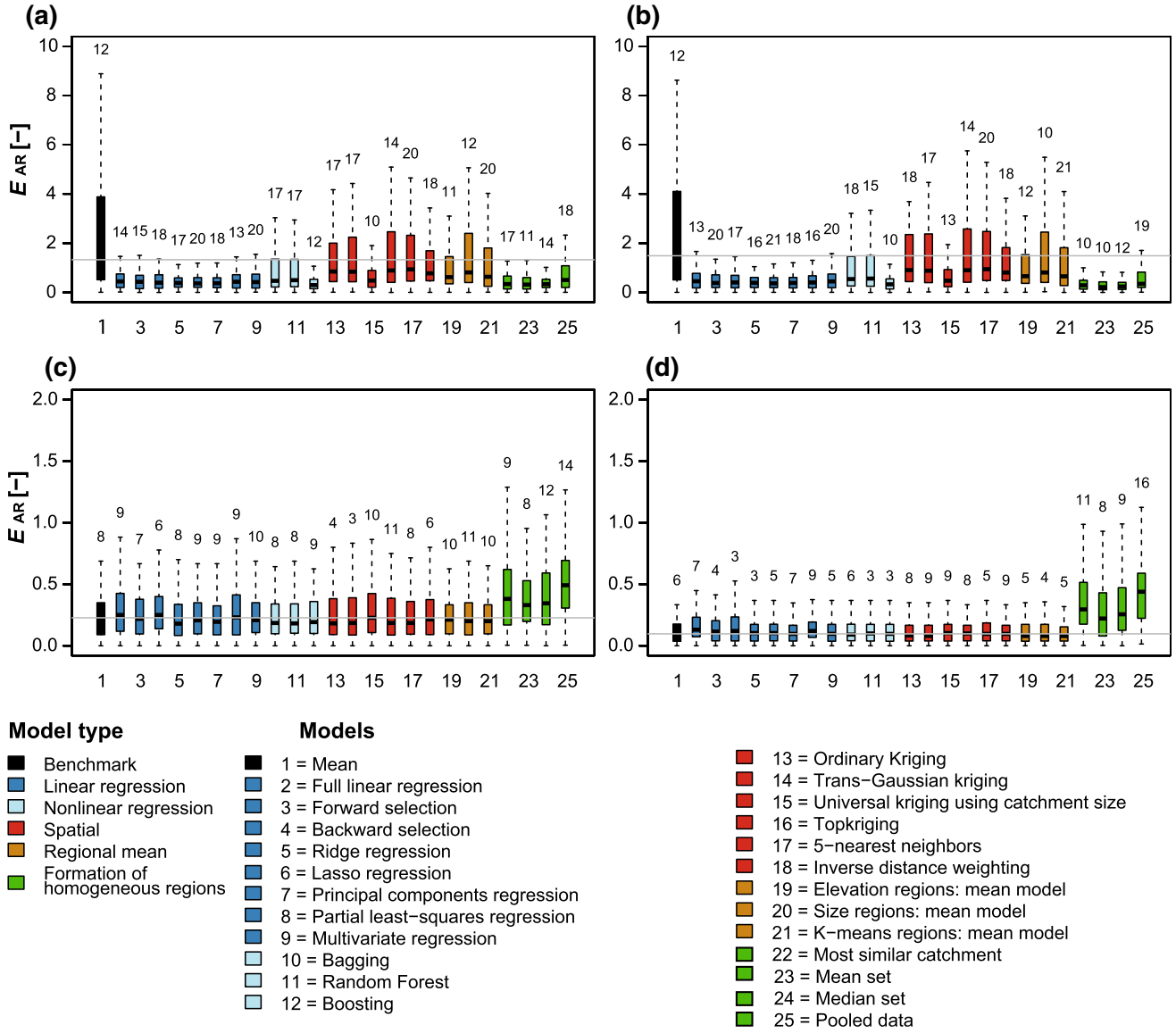



**Fig. 8** Predictive performance of the different regionalization methods tested for the different hydrograph characteristics **a** peak discharge ($Q_p$), **b** hydrograph volume ($V$), **c** time to peak ($t_p$), and **d** half-recession time ($t_{p05}$) provided as boxplots of the absolute relative error ($E_{AR}$) for the 163 catchments. The number of outliers (defined as those observations lying outside the quartile ± 1.5 times the interquartile range) is indicated by the numbers plotted above the boxplots

the magnitude of the event in terms of peak discharge and hydrograph volume as well as the baseflow component. However, the results also indicate that the hydrograph shape represented by the two parameters of the PDF is difficult to regionalize. The two components of the hydrograph, magnitude and shape, are addressed in turn.

*Hydrograph magnitude* Nonlinear regression methods showed a good performance with respect to the SDH parameters related to the magnitude of the flood hydrograph and also performed better than other methods looking at the four hydrograph characteristics ($Q_p$, $V$, $t_p$, and $t_{p05}$). Boosting performed best among the nonlinear regression techniques. In contrast to bagging and random forest, boosting not only reduces variance by combining several trees but also reduces bias thanks to the stagewise procedure in which successive trees use information from previous trees. Regionalization using boosted regression trees resulted in a median relative error of 50% over the different catchments for the peak discharges and 55% for hydrograph volumes computed for a return period of 100 years. This is comparable to studies by Petroselli and Grimaldi (2015) and Viviroli et al. (2009a).

Some of the linear regression methods performed as well as the nonlinear methods for those SDH parameters that are linked to the marginal distributions of peak discharges and hydrograph volumes. However, linear

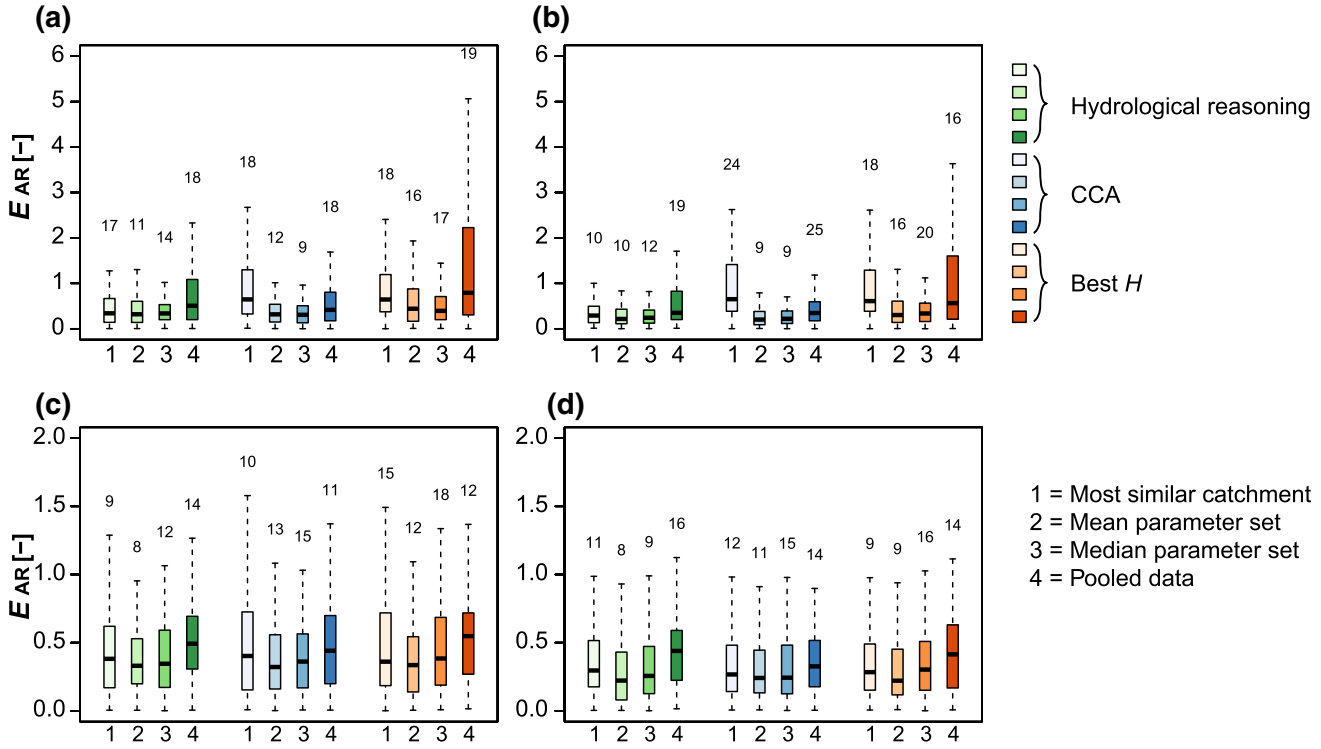



**Fig. 9** Predictive performance of the regionalization methods based on the formation of homogeneous regions in terms of boxplots of the absolute relative error ($E_{AR}$) for the hydrograph characteristics **a** peak discharge, **b** hydrograph volume, **c** time to peak and **d** half-recession time of the hydrograph. The strategies used to form homogeneous regions are: hydrological reasoning (green), CCA (blue), and best-*H* (red). The methods used to transfer the parameter set to the target catchment from its five most similar catchments are: (1) Transfer from most similar catchment, (2) Transfer of mean parameter set from five most similar catchments, (3) Transfer of median parameter set from five most similar catchments, and (4) Estimation of the parameter set based on the pooled data from the five most similar catchments

**Table 4** Catchment characteristics important for the prediction of the ten SDH parameters in ungauged catchments

| Parameters related to hydrograph | SDH parameter | Important catchment characteristics |
|---|---|---|
| Magnitude | $I_{BF}$ | Geology hard rock impermeable, X-coordinate, hydrogeology unconsolidated rock intermediate permeability |
| | GPD location | Catchment area, hydraulic topographic index standard deviation, length of main channel, relief energy |
| | GPD scale | Catchment area, maximum precipitation intensity over 24 h, hydraulic topographic index standard deviation |
| | GPD shape | West exposed surfaces, geology hard rock -pores, fissures, or karst, sunshine duration |
| | GEV location | Catchment area, hydraulic topographic index standard deviation, hydraulic conductivity skewness, relief energy |
| | GEV scale | Catchment area, hydraulic topographic index standard deviation, relief energy |
| | GEV shape | North exposed surfaces, maximum precipitation intensity, relief energy |
| | $\theta$ | South exposed surfaces, X-coordinate, hydrogeology hard rock impermeable, karstic rock, maximum precipitation during 24 h, average day of year |
| Shape | PDF location | Hydraulic topographic index skewness, forest, karstic rock |
| | PDF scale | Karstic rock, soil topographic index skewness, shape parameter 2 |

regression methods performed worse than nonlinear methods for the dependence between peak discharges and hydrograph volumes and the baseflow index. Yet, multiple regression methods have been found to be suitable for baseflow index regionalization by Haberlandt et al. (2001). Generally, some of the linear regression methods based on a reduction of the potential explanatory variables perform only slightly better than the full regression model. This might be related to the problem of finding appropriate explanatory variables.

Spatial methods could not compete with the regression methods when regionalizing the parameters of the marginal distributions of peak discharges and hydrograph volumes. The rather poor performance of kriging methods might be explained by the fact that the range of the empirical variogram was less than the mean distance between stations. However, they performed similarly well when regionalizing the dependence parameter and the baseflow index. The finding that regression-based methods generally perform better than spatial methods are in line with the recommendations made by the Centre for Eology and Hydrology (1999) but differ from observations made by Parajka et al. (2013), who analyzed 34 regionalization studies and concluded that regression-based methods are usually outperformed by spatial methods. He et al. (2011) found that spatial proximity was not helpful in all regionalization studies they reviewed because catchment characteristics can change abruptly in space. This is the case in Switzerland where even spatially close catchments can have quite different catchment characteristics. Spatial methods might perform better in regions with a high density of gauging stations (Merz 2006). In regions with fewer stations, such as in the alpine region, the uncertainty in the empirical variogram is expected to be larger (Blöschl 2006; Oudin et al. 2008). Among the spatial methods, universal kriging using catchment area as an explanatory variable clearly performed best when predicting the event magnitude. This is not surprising since we have seen that catchment area is one of the most important explanatory variables in the data set. Removing the scale effect of different catchment sizes is therefore very important. The other spatial methods did not differ significantly in their predictive performance. We did not find that top-kriging outperformed ordinary kriging as was found by Skoien et al. (2006). This is likely because spatial proximity does not necessarily entail similarity in hydrological behavior (Hrachowitz et al. 2013). In addition, the poor performance of top-kriging might also be related to the fact that not all the SDH parameters were well suited for the application of the top-kriging approach. The approach was developed for variables that accumulate along the stream network (Skoien et al. 2006), which is not necessarily the case for parameters of a statistical distribution.

The regional mean models led to a slight improvement compared to the global mean model (benchmark), which was also found by Sauquet and Catalogne (2011), but did not perform as well as regression methods for the regionalization of the flood magnitude. The methods transferring the entire parameter set from similar catchments to the target catchments led to good results in the prediction of peak discharges and hydrograph volumes. However, it was difficult to define regions that were homogeneous in terms of both catchment and hydrologic characteristics (*L*-moments of peak discharges) (Shu and Ouarda 2008) using all of the three strategies applied.

*Hydrograph shape* The two SDH parameters related to the hydrograph shape (PDF location and PDF scale) could not be satisfactorily regionalized. This means that no regionalization method could be found that outperformed the benchmark model. This finding was supported by the prediction error for the two hydrograph characteristics time to peak and half-recession time. This is in line with findings by Cipriani et al. (2012) who found that flood durations were more difficult to regionalize than flood quantiles. The reason for the poor performance in hydrograph shape regionalization might be that a part of the variability in flood shapes within a catchment was reduced when defining the representative normalized hydrograph (RNH) via the median hydrograph. As a consequence, RNHs across catchments were quite similar. Therefore, it is sufficient to work with one mean hydrograph shape across all catchments. Eventually, the nonparametric "distance-based" approach proposed by Ganora et al. (2009) for the regionalization of flow duration curves could be adjusted and successfully used for the regionalization of hydrograph shapes.

While the errors of the hydrograph volumes and peak discharges were correlated, errors in the hydrograph shape did not imply errors in the magnitude of the event. This means that even if an accurate prediction of the shape is not possible, the magnitude of the event can still be predicted accurately.

### 5.2 Best regionalization model

Hydrograph magnitudes were best regionalized using the nonlinear regression technique boosted regression trees while the mean model (benchmark) is the best model to regionalize the hydrograph shape. This difference shows that no universal regionalization method can be found (Razavi and Coulibaly 2013) for the regionalization of SDHs. One could take this into account by combining the two regionalization methods. Boosted regression trees could be applied to regionalize the SDH parameters related to the hydrograph magnitude while the mean model could be used for the regionalization of the SDH parameters

related to the hydrograph shape. Such a combined approach might lead to more accurate flood estimates in ungauged catchments than the use of one single method.

### 5.3 Important catchment characteristics for prediction

Different variables were important in fitting the different types of linear and nonlinear regression models for the ten SDH parameters. We again look first at the SDH parameters related to the hydrograph magnitude and then at the two SDH parameters related to the hydrograph shape.

*Hydrograph magnitude*

- Parameter baseflow: We found that geological and hydrogeological features were important for predicting the baseflow index, which is one of the most important low flow indices. Geological features have also been found to be important for the prediction of the baseflow index in other regions such as the Mediterranean (Longobardi and Villani 2008).
- Parameters marginal distributions: Similar catchment characteristics were important for the prediction of the location and scale parameters of the marginal distributions of peak discharges and hydrograph volumes. Namely, these were: catchment area, length of the main channel, maximum precipitation intensity, relief energy, and a parameter related to the topographic hydraulic index. Similar catchment characteristics were also found to be important in the prediction of flood quantiles by Ahn and Palmer (2016), Haddad and Rahman (2012) and Rahman et al. (2017). This is not surprising, because the magnitude of a flood is expected to increase with catchment size and the length of the main channel.
- Parameter dependence: Location in space and the exposition of surfaces was important for the prediction of the SDH parameter modeling the dependence between peak discharges and hydrograph volumes.

*Hydrograph shape* We found that the presence or absence of karstic rock is important for the prediction of the SDH parameters characterizing the hydrograph shape. However, no suitable regionalization method can be found for the SDH parameters related to the hydrograph shape, which means that it is a very weak explanatory variable.

### 5.4 Limitations

The regionalization of SDH parameters to ungauged catchments is hampered by three main factors: the lack of suitable catchment characteristics, the presence of uncertainties in both the computation of SDHs based on observed runoff and their regionalization, and by the fact that one hydrograph shape within a catchment is not representative of hydrograph shape variability.

*Lack of suitable catchment characteristics* The nonlinear regression methods take into account the fact that hydrological relationships are unlikely to be linear in nature (Aziz et al. 2015; Parajka et al. 2005) but they are unable to uncover underlying physical laws (He et al. 2011) like all regression based analyses. In addition, they cannot overcome the problem that the available catchment characteristics do not sufficiently explain the hydrological behavior and the flood generating processes of a catchment (Ali et al. 2012; Salinas et al. 2013). We need catchment characteristics that better represent the hydrological behavior of a catchment and an improved understanding of physical processes in a catchment to make a further step towards better prediction (He et al. 2011; Merz and Blöschl 2003; Oudin et al. 2010). However, the choice of relevant catchment characteristics is likely to be region and/or context dependent (Ali et al. 2012).

*Uncertainties* The quality of the regionalized values depends on two factors (Merz 2006): first, on the uncertainty of SDH estimates for gauged catchments; second, on the uncertainty inherent in the regionalization of those parameters. The uncertainty of the SDH parameters in the gauged catchments results from the length of the observation record, different model choices, and parameter estimation during the SDH construction process. The regionalization uncertainty can be considerable (Petroselli and Grimaldi 2015) and is among other sources of uncertainty related to the choice of a particular regionalization model out of a selection of coherent, sensible approaches. This source of uncertainty can be quantified thanks to this comprehensive comparison of regionalization methods. As an ensemble, the hydrographs obtained by the different regionalization methods give an idea of the variability introduced due to the choice of a regionalization model (McIntyre et al. 2005). For a more detailed analysis of both SDH construction and regionalization uncertainty, the reader is referred to the simulation study by Brunner et al. (2017a).

*Misrepresentation of hydrograph shape variability* Representing the hydrograph shape by just one catchment specific shape and regionalizing it with a mean model neglects the variability in flood shapes within a catchment. The hydrograph shape is only partly catchment specific and can differ significantly between different flood types, such as long-rain floods, short-rain floods, flash floods, or rain-on-snow floods (Sikorska et al. 2015).

### 5.5 Perspectives

The limitations mentioned above could be addressed as follows:

*Identification of suitable catchment characteristics* A better understanding of the influence of catchment characteristics on the flood response of a catchment might help to find suitable catchment characteristics for the regionalization of SDHs. However, the flood response is not only dependent on catchment characteristics but also on the particular weather pattern and antecedent wetness conditions that trigger the event. Nied et al. (2014) found that flood favoring hydro-meteorological patterns vary between seasons. They therefore claimed that flood frequency analysis should try to describe flood occurrence in dependence of hydro-meteorological patterns.

*Uncertainty assessment* To overcome the uncertainty related to the choice of the regionalization method, a combination of several methods as suggested by Merz (2006) might be appropriate. A combination of different methods, tested in this study, provides information on the variability of the design flood estimated for an ungauged catchment and could reduce the uncertainty related to a single estimate (Deutshe Vereinigung für Wasserwirtschaft Abwasser und Abfall 2012). Still, the uncertainty coming from the regionalization is not the only uncertainty source to be considered. The total uncertainty of an SDH in an ungauged catchment is also affected by the uncertainty of the data used for regionalization, i.e., the SDHs constructed based on observed runoff data.

*Better representation of hydrograph shape variability* The variability caused by different flood types could be accounted for by considering flood-type specific hydrograph shapes (Brunner et al. 2017b). These would take into account the heterogeneity of climatic inputs that make the prediction of a basin response difficult (Sivapalan 2003). However, we found that such flood-type specific hydrograph shapes are not very useful for regionalization since identifying groups of catchments with similar flood type patterns is difficult. We should therefore think about alternatives for the representation of the hydrograph shape variability within a catchment. Allowing for the representation of more than one shape type per catchment might reduce the uncertainty of the hydrograph shape estimates and therefore facilitate their regionalization.

## 6 Conclusions

We tested 24 methods, among them three nonlinear regression techniques not commonly applied in regionalization studies, for the regionalization of synthetic design hydrographs that provide information on the hydrograph peak and on the whole hydrograph including hydrograph volume, time to peak, and half-recession time. The extensive method comparison allowed for the identification of the most suitable regionalization method for synthetic design hydrographs. We showed that the regionalization of the design flood magnitude is possible, while the shape of the design flood remains difficult to regionalize. The parameters related to the hydrograph magnitude were best regionalized using the nonlinear approach of boosted regression trees. On the other hand, the mean model is sufficient to regionalize the parameters related to the hydrograph shape. Nonlinear regression methods were found to result in a better predictive performance than linear regression methods, spatial methods, and several other methods based on the formation of homogeneous regions. Spatial proximity alone was not able to explain the differences in design floods for different catchments. These differences can be better explained by different physiographical and climatological catchment characteristics. Among them, catchment area, length of the main channel, and relief energy were most important for the regionalization of SDH parameters related to the magnitude of the event. The relationship between these characteristics and the characteristics of the design flood are not always linear and are therefore best explained using nonlinear regression models. The regionalized design floods remain uncertain despite a comprehensive comparison of different regionalization methods. Uncertainties come from both the estimation of synthetic design hydrographs in gauged catchments and the regionalization of these hydrographs to ungauged catchments. These uncertainties need to be properly assessed in a simulation study to increase the reliability of the design hydrograph estimates. We recommend to use tree-based models such as bagging, random forest, and boosting in future regionalization studies. They can be especially useful when the relationship between a hydrological variable and different potential explanatory variables is complex.

**Acknowledgements** We thank the Federal Office for the Environment (FOEN) for funding the project (contract 13.0028.KP/M285-0623) and for providing runoff measurement data. We also thank MeteoSwiss for providing precipitation data. The data used in this study is available upon order from the FOEN and MeteoSwiss. For the hydrological data of the federal stations, the order form under http://www.bafu.admin.ch/wasser/13462/13494/15076/index.html?lang=de can be used. The hydrological data of the cantonal stations can be ordered from the respective cantons. The meteorological data can be ordered via https://shop.meteoswiss.ch/index.html. We thank the associate editor and the four reviewers for their constructive and detailed comments.

## Appendix: List of stations used in this regionalization study

See Table 5.

**Table 5** List of stations used in this regionalization study, a summary of their catchment characteristics, and their locally estimated SDH parameters (last ten columns)

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner | $I_{BF}$ | PDF location | PDF scale | GPD location | GPD scale | GPD shape | GEV location | GEV scale | GEV shape | $\theta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Aabach–Mönchaltorf | 46 | 440 | 521 | 0 | 34 | ZH | 0.41 | − 1.31 | 0.77 | 6 | 8.56 | − 0.15 | 0.57 | 0.21 | 0.11 | 2.07 |
| 2 | Aach–Salmsach | 49 | 406 | 480 | 0 | 40 | FOEN | 0.39 | − 1.08 | 0.61 | 4.7 | 6.39 | − 0.15 | 0.49 | 0.17 | 0.17 | 1.97 |
| 3 | Aire–Confignon | 57 | 398 | 454 | 0 | 23 | GE | 0.3 | − 1.1 | 0.44 | 6.9 | 15.89 | − 0.25 | 0.75 | 0.22 | 0.15 | 1.91 |
| 4 | Allenbach–Adelboden | 29 | 1297 | 1856 | 0 | 40 | FOEN | 0.36 | − 1.32 | 0.79 | 3.6 | 7.42 | 0.07 | 0.32 | 0.19 | 0.13 | 1.27 |
| 5 | Alpbach–Erstfeld | 21 | 1019 | 2200 | 28 | 41 | FOEN | 0.47 | − 1.27 | 0.7 | 5.8 | 5.11 | 0.06 | 0.48 | 0.15 | 0.03 | 1.57 |
| 6 | Alp–Einsiedeln | 46 | 840 | 1155 | 0 | 23 | FOEN | 0.41 | − 1.18 | 0.56 | 14.5 | 35.32 | − 0.29 | 1.19 | 0.46 | 0.02 | 1.25 |
| 7 | Altbach–Bassersdorf | 13 | 470 | 549 | 0 | 37 | ZH | 0.42 | − 0.77 | 0.42 | 1 | 1.41 | − 0.04 | 0.1 | 0.04 | 0.13 | 1.97 |
| 8 | Arbogne–Avenches | 70 | 435 | 597 | 0 | 20 | VD | 0.58 | − 0.97 | 0.5 | 3.1 | 3.83 | − 0.09 | 0.24 | 0.1 | 0.3 | 2.24 |
| 9 | Areuse–Boudry | 377 | 444 | 1060 | 0 | 31 | FOEN | 0.35 | − 0.88 | 0.64 | 47.6 | 17.59 | − 0.07 | 6.07 | 1.72 | 0.08 | 3.41 |
| 10 | Arnon–Grandson | 83 | 434 | 942 | 0 | 20 | VD | 0.46 | − 0.99 | 0.55 | 12.2 | 5.16 | − 0.26 | 1.22 | 0.41 | 0.01 | 1.99 |
| 11 | Aubonne–Allaman | 91 | 390 | 890 | 0 | 35 | FOEN | 0.2 | − 0.71 | 0.67 | 16.3 | 8.72 | − 0.01 | 2.36 | 0.78 | 0.02 | 2.53 |
| 12 | Augstbach–Balsthal | 64 | 485 | 801 | 0 | 20 | SO | 0.36 | − 1.04 | 0.59 | 5.4 | 6.35 | 0.19 | 0.58 | 0.23 | 0.19 | 1.71 |
| 13 | Biber–Biberbrugg | 32 | 825 | 1009 | 0 | 25 | FOEN | 0.33 | − 0.82 | 0.55 | 7.7 | 16.51 | − 0.55 | 0.74 | 0.3 | 0.13 | 1.56 |
| 14 | Bibere–Kerzers | 50 | 443 | 540 | 0 | 34 | FOEN | 0.61 | − 0.95 | 0.66 | 2.5 | 2.98 | 0.07 | 0.16 | 0.07 | 0.24 | 2.13 |
| 15 | Birse–Court | 92 | 663 | 925 | 0 | 20 | BE | 0.53 | − 1.06 | 0.64 | 7.1 | 8.45 | − 0.18 | 0.55 | 0.22 | 0.01 | 1.65 |
| 16 | Birse–Moutier | 183 | 519 | 930 | 0 | 40 | FOEN | 0.58 | − 0.91 | 0.49 | 12.9 | 11.05 | − 0.13 | 0.94 | 0.35 | 0.14 | 1.92 |
| 17 | Birse–Soyhières | 590 | 395 | 810 | 0 | 28 | FOEN | 0.47 | − 1 | 0.62 | 43.3 | 25.55 | 0.04 | 4.48 | 1.37 | 0.27 | 2.21 |
| 18 | Birsig–Binningen | 75 | 281 | 434 | 0 | 35 | BL | 0.5 | − 1.11 | 0.66 | 3.6 | 4.26 | 0.05 | 0.34 | 0.14 | 0.2 | 1.77 |
| 19 | Birs–Münchenstein | 911 | 268 | 726 | 0 | 40 | FOEN | 0.47 | − 1.02 | 0.56 | 66.8 | 31.15 | 0.02 | 6.84 | 1.98 | 0.2 | 2.65 |
| 20 | Breggia–Chiasso | 47 | 255 | 927 | 0 | 40 | FOEN | 0.23 | − 1.08 | 0.51 | 12.9 | 27.95 | − 0.11 | 1.69 | 0.61 | 0.03 | 1.81 |
| 21 | Brinaz–Yverdon–les–Bains | 14 | 434 | 542 | 0 | 20 | VD | 0.45 | − 1.06 | 0.52 | 1.5 | 3.55 | − 0.12 | 0.15 | 0.06 | 0.11 | 2.73 |
| 22 | Broye–Payerne | 392 | 441 | 710 | 0 | 40 | FOEN | 0.43 | − 0.89 | 0.48 | 48.3 | 72.08 | − 0.44 | 4.63 | 1.63 | 0.06 | 1.82 |
| 23 | Bruggbach–Gipf/Oberfrick | 45 | 356 | 575 | 0 | 35 | AG | 0.38 | − 1.09 | 0.64 | 3.6 | 6.51 | − 0.16 | 0.4 | 0.16 | 0.13 | 1.53 |
| 24 | Bünz–Muri (Hasli) | 15 | 448 | 613 | 0 | 30 | AG | 0.9 | − 0.74 | 0.44 | 1.3 | 2.33 | 0.14 | 0.1 | 0.04 | 0.2 | 1.76 |
| 25 | Bünz–Othmarsingen | 111 | 390 | 533 | 0 | 37 | AG | 0.56 | − 0.61 | 0.4 | 6.8 | 6.87 | 0.02 | 0.49 | 0.19 | 0.36 | 2.12 |
| 26 | Bünz–Wohlen | 53 | 421 | 555 | 0 | 34 | AG | 0.53 | − 0.92 | 0.46 | 3.3 | 5.45 | − 0.07 | 0.26 | 0.12 | 0.4 | 2.23 |
| 27 | Buuserbach–Maisprach | 11 | 367 | 529 | 0 | 36 | BL | 0.65 | − 1.03 | 0.62 | 0.6 | 0.64 | 0.01 | 0.04 | 0.02 | 0.15 | 1.35 |
| 28 | Cassarate–Pregassona | 74 | 291 | 990 | 0 | 40 | FOEN | 0.4 | − 0.76 | 0.62 | 9.5 | 15.69 | 0 | 1.21 | 0.53 | 0.13 | 1.65 |
| 29 | Chämtnerbach− Wetzikon | 13 | 560 | 760 | 0 | 29 | ZH | 0.4 | − 0.61 | 0.29 | 1.9 | 3.62 | − 0.06 | 0.17 | 0.07 | 0.18 | 1.81 |
| 30 | Chandon–Avenches | 39 | 432 | 571 | 0 | 20 | VD | 0.66 | − 0.68 | 0.36 | 1.8 | 3.64 | 0.07 | 0.09 | 0.04 | 0.54 | 2.34 |
| 31 | Chli Schliere–Alpnach | 22 | 453 | 1370 | 0 | 36 | FOEN | 0.38 | − 0.8 | 0.55 | 4.4 | 8.83 | 0.04 | 0.37 | 0.14 | 0.32 | 1.94 |
| 32 | Chrebsbach–St. Margarethen | 14 | 503 | 581 | 0 | 21 | TG | 0.39 | − 1.3 | 0.74 | 1.2 | 3.88 | − 0.21 | 0.12 | 0.05 | 0.21 | 1.92 |
| 33 | Diegterbach–Diegten | 13 | 509 | 746 | 0 | 31 | BL | 0.43 | − 1.16 | 0.67 | 1.4 | 2.28 | 0.06 | 0.11 | 0.05 | 0.23 | 1.35 |

**Table 5** (continued)

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner | $I_{BF}$ | PDF location | PDF scale | GPD location | GPD scale | GPD shape | GEV location | GEV scale | GEV shape | $\theta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 34 | Diegterbach–Sissach | 33 | 372 | 614 | 0 | 36 | BL | 0.32 | − 1.15 | 0.58 | 2.7 | 5.11 | − 0.13 | 0.29 | 0.11 | 0.29 | 1.58 |
| 35 | Dorfbach–Allschwil | 11 | 281 | 360 | 0 | 30 | BL | 0.59 | − 1.16 | 0.67 | 0.3 | 0.48 | 0.43 | 0.02 | 0.01 | 0.43 | 2.75 |
| 36 | Drize–Lancy | 23 | 392 | 528 | 0 | 25 | GE | 0.32 | − 1.33 | 0.65 | 2 | 4.33 | − 0.23 | 0.22 | 0.08 | 0.17 | 2.09 |
| 37 | Dünnern–Olten | 196 | 400 | 750 | 0 | 36 | FOEN | 0.37 | − 1.15 | 0.53 | 16.6 | 20.52 | − 0.11 | 1.8 | 0.7 | 0.16 | 1.9 |
| 38 | Eibach–Gelterkinden | 27 | 405 | 627 | 0 | 36 | BL | 0.23 | − 0.99 | 0.67 | 2.4 | 2.78 | 0.12 | 0.3 | 0.11 | 0.17 | 1.66 |
| 39 | Eibach–Zeglingen | 13 | 517 | 725 | 0 | 30 | BL | 0.35 | − 1.1 | 0.68 | 1.2 | 1.51 | 0.21 | 0.13 | 0.05 | 0.13 | 1.55 |
| 40 | Emme–Eggiwil | 124 | 745 | 1189 | 0 | 39 | FOEN | 0.29 | − 0.96 | 0.49 | 27.1 | 44.22 | − 0.21 | 2.68 | 0.99 | 0.08 | 1.53 |
| 41 | Emme–Wiler | 939 | 458 | 860 | 0 | 40 | FOEN | 0.54 | − 0.64 | 0.38 | 84.2 | 80.96 | − 0.01 | 6.01 | 2.06 | 0.21 | 1.66 |
| 42 | Ergolz–Itingen | 141 | 350 | 593 | 0 | 33 | BL | 0.34 | − 1.02 | 0.6 | 10 | 10.26 | 0 | 1.14 | 0.39 | 0.23 | 1.76 |
| 43 | Ergolz–Liestal | 261 | 305 | 590 | 0 | 40 | FOEN | 0.34 | − 1 | 0.66 | 19.9 | 20.27 | 0 | 2.25 | 0.78 | 0.23 | 2.16 |
| 44 | Ergolz–Ormalingen | 30 | 410 | 585 | 0 | 36 | BL | 0.24 | − 1 | 0.66 | 2.4 | 3.55 | − 0.01 | 0.33 | 0.13 | 0.21 | 1.93 |
| 45 | Etzgerbach–Etzgen | 25 | 308 | 478 | 0 | 34 | AG | 0.45 | − 0.9 | 0.45 | 2.4 | 4.07 | − 0.1 | 0.19 | 0.06 | 0.19 | 1.46 |
| 46 | Eulach–Wülflingen | 73 | 410 | 532 | 0 | 43 | ZH | 0.35 | − 1.22 | 0.57 | 4.1 | 9.54 | − 0.05 | 0.42 | 0.2 | 0.21 | 1.46 |
| 47 | Fisibach–Fisibach | 15 | 379 | 516 | 0 | 31 | AG | 0.65 | − 0.63 | 0.56 | 0.6 | 0.72 | 0.09 | 0.03 | 0.01 | 0.3 | 1.85 |
| 48 | Flon–Oron–la− Ville | 16 | 609 | 812 | 0 | 20 | VD | 0.25 | − 1.19 | 0.76 | 2.2 | 3.88 | − 0.18 | 0.25 | 0.09 | 0.17 | 2.14 |
| 49 | Furtbach–Würenlos | 39 | 410 | 482 | 0 | 36 | ZH | 0.57 | − 1.08 | 0.63 | 2.8 | 2.11 | − 0.05 | 0.22 | 0.08 | 0.13 | 1.99 |
| 50 | Geisslibach–Furtmüli | 20 | 415 | 474 | 0 | 24 | ZH | 0.68 | − 0.48 | 0.31 | 0.6 | 0.63 | 0.01 | 0.04 | 0.02 | 0.24 | 1.59 |
| 51 | Glatt–Herisau | 16 | 679 | 840 | 0 | 40 | FOEN | 0.45 | − 0.93 | 0.51 | 3.3 | 5.84 | 0.33 | 0.26 | 0.11 | 0.04 | 1.72 |
| 52 | Goldach–Goldach | 50 | 399 | 833 | 0 | 23 | FOEN | 0.42 | − 1.28 | 0.82 | 8.2 | 13.82 | 0.05 | 0.76 | 0.33 | 0.2 | 1.71 |
| 53 | Goneri–Oberwald | 40 | 1385 | 2377 | 14 | 23 | FOEN | 0.55 | − 0.85 | 0.43 | 7.2 | 9.82 | − 0.02 | 0.64 | 0.24 | 0.21 | 1.85 |
| 54 | Grenet (amont)–Pigeon | 19 | 680 | 748 | 0 | 20 | VD | 0.27 | − 1.26 | 0.54 | 4.4 | 8.53 | − 0.42 | 0.45 | 0.14 | 0.02 | 1.23 |
| 55 | Grossbach–Gross | 11 | 900 | 1235 | 0 | 40 | FOEN | 0.36 | − 0.57 | 0.36 | 2.5 | 4.23 | 0.07 | 0.2 | 0.08 | 0.07 | 1.28 |
| 56 | Grosstalbach–Isenthal | 44 | 767 | 1820 | 9 | 40 | FOEN | 0.59 | − 1.19 | 0.8 | 5 | 3.26 | 0.33 | 0.39 | 0.15 | 0.13 | 1.73 |
| 57 | Gürbe–Belp | 117 | 511 | 837 | 0 | 40 | FOEN | 0.55 | − 1.01 | 0.55 | 14.9 | 9.54 | − 0.21 | 0.85 | 0.27 | 0.07 | 1.36 |
| 58 | Gürbe–Burgistein | 54 | 568 | 1044 | 0 | 28 | FOEN | 0.43 | − 1.2 | 0.69 | 7 | 8.87 | − 0.01 | 0.61 | 0.2 | 0.17 | 1.44 |
| 59 | Haselbach–Maschwanden | 20 | 390 | 495 | 0 | 37 | ZH | 0.51 | − 0.8 | 0.63 | 1.7 | 2.72 | − 0.17 | 0.14 | 0.06 | 0.26 | 2.34 |
| 60 | Hintere Frenke–Bubendorf | 38 | 352 | 603 | 0 | 30 | BL | 0.33 | − 1.08 | 0.68 | 3.1 | 3.44 | 0.05 | 0.37 | 0.13 | 0.14 | 1.73 |
| 61 | Hintere Frenke–Reigoldswil | 15 | 489 | 742 | 0 | 32 | BL | 0.47 | − 1.1 | 0.72 | 1.6 | 1.2 | 0.2 | 0.15 | 0.05 | 0.18 | 1.79 |
| 62 | Hinterrhein–Hinterrhein | 54 | 1584 | 2360 | 17 | 35 | FOEN | 0.58 | − 0.71 | 0.58 | 16.5 | 20.93 | − 0.04 | 1.09 | 0.42 | 0.22 | 1.82 |
| 63 | Holzbach–Villmergen | 24 | 416 | 590 | 0 | 34 | AG | 0.52 | − 1.04 | 0.7 | 1.4 | 1.93 | − 0.06 | 0.11 | 0.05 | 0.34 | 1.85 |
| 64 | Homburgerbach–Thürnen | 30 | 387 | 615 | 0 | 36 | BL | 0.2 | − 0.94 | 0.6 | 2.2 | 2.3 | 0.09 | 0.29 | 0.09 | 0.14 | 1.59 |
| 65 | Ilfis–Langnau | 188 | 685 | 1051 | 0 | 25 | FOEN | 0.48 | − 0.87 | 0.38 | 25.6 | 44.32 | − 0.01 | 2.16 | 0.78 | 0.27 | 1.8 |
| 66 | Jona–Pilgersteg | 24 | 560 | 818 | 0 | 44 | ZH | 0.46 | − 1.01 | 0.56 | 5.4 | 7.01 | − 0.09 | 0.47 | 0.17 | 0.1 | 1.69 |
| 67 | Jona–Rüti | 58 | 450 | 669 | 0 | 20 | ZH | 0.43 | − 0.72 | 0.44 | 11.1 | 13.67 | − 0.15 | 0.96 | 0.39 | 0.17 | 2.03 |

**Table 5** (continued)

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner | $I_{BF}$ | PDF location | PDF scale | GPD location | GPD scale | GPD shape | GEV location | GEV scale | GEV shape | $\theta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 68 | Jonen–Zwillikon | 39 | 460 | 605 | 0 | 27 | ZH | 0.42 | − 0.98 | 0.56 | 3.3 | 4.64 | 0.07 | 0.32 | 0.13 | 0.26 | 2.06 |
| 69 | Kaisterbach–Kaisten | 12 | 321 | 464 | 0 | 34 | AG | 0.51 | − 0.95 | 0.5 | 0.8 | 1.59 | 0.08 | 0.06 | 0.02 | 0.23 | 1.51 |
| 70 | Kander–Hondrich | 520 | 650 | 1900 | 8 | 33 | FOEN | 0.72 | − 0.82 | 0.51 | 47.9 | 23.93 | − 0.06 | 2.67 | 0.94 | 0.08 | 1.94 |
| 71 | Kempt–Fehraltorf | 24 | 520 | 645 | 0 | 23 | ZH | 0.47 | − 1.04 | 0.49 | 1.9 | 2.98 | 0.07 | 0.17 | 0.08 | 0.25 | 2.4 |
| 72 | Kempt–Winterthur | 60 | 450 | 588 | 0 | 33 | ZH | 0.5 | − 1.1 | 0.51 | 6.3 | 5.75 | 0.05 | 0.56 | 0.22 | 0.18 | 2.14 |
| 73 | Kleine Emme–Littau | 477 | 431 | 1050 | 0 | 36 | FOEN | 0.45 | − 0.67 | 0.41 | 84.5 | 119.79 | − 0.18 | 6.63 | 2.25 | 0.17 | 1.75 |
| 74 | Kleine Emmev–Werthenstein | 311 | 540 | 1173 | 0 | 30 | FOEN | 0.45 | − 0.97 | 0.62 | 56.7 | 74.78 | − 0.14 | 4.55 | 1.61 | 0.16 | 1.77 |
| 75 | Köllikerbach–Kölliken | 10 | 423 | 488 | 0 | 31 | AG | 0.47 | − 0.79 | 0.43 | 1.2 | 2.99 | − 0.19 | 0.08 | 0.04 | 0.17 | 1.38 |
| 76 | Langeten–Huttwil | 60 | 597 | 766 | 0 | 40 | FOEN | 0.6 | − 1 | 0.57 | 4 | 5.51 | 0.11 | 0.24 | 0.1 | 0.23 | 1.58 |
| 77 | Langeten–Lotzwil | 115 | 500 | 713 | 0 | 20 | BE | 0.66 | − 1.05 | 0.75 | 5.8 | 1.46 | − 0.11 | 0.31 | 0.11 | − 0.1 | 1.35 |
| 78 | Lonza–Blatten | 78 | 1520 | 2630 | 37 | 40 | FOEN | 0.43 | − 1 | 0.92 | 11.6 | 8.83 | − 0.15 | 1.17 | 0.45 | − 0.08 | 1.11 |
| 79 | Louibach–Saanen | 62 | 1085 | 1875 | 6 | 20 | BE | 0.64 | − 0.82 | 0.51 | 8.4 | 4.98 | − 0.22 | 0.48 | 0.18 | 0.09 | 1.45 |
| 80 | Luthern–Nebikon | 108 | 494 | 740 | 0 | 26 | FOEN | 0.49 | − 0.89 | 0.51 | 6.5 | 9.9 | − 0.07 | 0.58 | 0.21 | 0.24 | 1.93 |
| 81 | Lyssbach–Lyss | 50 | 444 | 574 | 0 | 22 | BE | 0.65 | − 0.6 | 0.31 | 3.1 | 6.45 | − 0.26 | 0.16 | 0.06 | 0.33 | 1.98 |
| 82 | Lyssbach–Schüpfen | 23 | 505 | 616 | 0 | 22 | BE | 0.65 | − 0.31 | 0.31 | 1.1 | 2.32 | − 0.11 | 0.06 | 0.02 | 0.43 | 1.94 |
| 83 | Magdenerbach–Rheinfelden | 33 | 300 | 483 | 0 | 31 | AG | 0.46 | − 1.05 | 0.96 | 1.8 | 2.64 | − 0.11 | 0.18 | 0.07 | 0.17 | 1.98 |
| 84 | Magliasina–Magliaso | 34 | 295 | 920 | 0 | 34 | FOEN | 0.3 | − 0.95 | 0.56 | 6 | 7.92 | 0.18 | 0.77 | 0.33 | 0.2 | 1.99 |
| 85 | Marchbach–Oberwil | 27 | 296 | 462 | 0 | 34 | BL | 0.58 | − 1.1 | 0.99 | 1.2 | 2.12 | − 0.01 | 0.08 | 0.04 | 0.33 | 1.57 |
| 86 | Mederbach–Marthalen | 26 | 375 | 439 | 0 | 46 | ZH | 0.52 | − 0.99 | 0.63 | 0.6 | 0.84 | 0.03 | 0.06 | 0.03 | 0.24 | 1.62 |
| 87 | Mederbach–Niederwiesen | 30 | 368 | 425 | 0 | 30 | ZH | 0.55 | − 1.01 | 0.66 | 0.6 | 0.6 | − 0.22 | 0.05 | 0.03 | 0.12 | 1.17 |
| 88 | Mentue–Yvonand | 105 | 449 | 679 | 0 | 40 | FOEN | 0.48 | − 0.68 | 0.41 | 9.3 | 14.69 | − 0.25 | 0.81 | 0.29 | 0.17 | 1.71 |
| 89 | Minster–Euthal | 59 | 894 | 1351 | 0 | 40 | FOEN | 0.36 | − 0.41 | 0.34 | 18.5 | 43.01 | − 0.22 | 1.58 | 0.61 | 0.09 | 1.21 |
| 90 | Möhlinbach–Zeiningen | 27 | 338 | 514 | 0 | 32 | AG | 0.41 | − 1.19 | 0.81 | 1.6 | 1.92 | − 0.1 | 0.16 | 0.06 | 0.11 | 1.83 |
| 91 | Murg–Frauenfeld | 212 | 390 | 580 | 0 | 40 | FOEN | 0.49 | − 1 | 0.62 | 18.4 | 22.27 | − 0.04 | 1.77 | 0.69 | 0.12 | 2.41 |
| 92 | Murg–Murgenthal | 207 | 419 | 637 | 0 | 34 | FOEN | 0.65 | − 1.02 | 0.73 | 11.2 | 7.68 | − 0.1 | 0.61 | 0.22 | 0.13 | 1.51 |
| 93 | Murg–Wängi | 79 | 466 | 650 | 0 | 40 | FOEN | 0.44 | − 0.85 | 0.41 | 8.6 | 8.52 | − 0.11 | 0.83 | 0.3 | 0.02 | 2.1 |
| 94 | Näfbach–Neftenbach | 38 | 394 | 464 | 0 | 22 | ZH | 0.56 | − 0.85 | 0.44 | 2 | 2.44 | 0.08 | 0.16 | 0.07 | 0.23 | 1.97 |
| 95 | Necker–Mogelsberg | 88 | 606 | 959 | 0 | 40 | FOEN | 0.35 | − 0.62 | 0.4 | 20.3 | 26.47 | 0.14 | 1.95 | 0.72 | 0.18 | 1.82 |
| 96 | Nozon–Pré Chaillet | 45 | 440 | 882 | 0 | 21 | VD | 0.28 | − 0.96 | 0.61 | 3.1 | 1.48 | − 0.04 | 0.37 | 0.09 | 0.24 | 2.6 |
| 97 | Önz–Heimenhusen | 84 | 440 | 583 | 0 | 20 | BE | 0.73 | − 0.84 | 0.51 | 4 | 3.02 | − 0.21 | 0.17 | 0.06 | 0.33 | 1.7 |
| 98 | Orbe–Le Sentier | 96 | 1010 | 1210 | 0 | 21 | VD | 0.34 | − 0.9 | 0.56 | 9.8 | 3.6 | − 0.07 | 1.16 | 0.31 | 0 | 2.37 |
| 99 | Orisbach–Liestal | 21 | 315 | 515 | 0 | 33 | BL | 0.2 | − 0.91 | 0.76 | 1.4 | 2.07 | 0.09 | 0.21 | 0.08 | 0.4 | 2.09 |
| 100 | Ova dal Fuorn–Zernez | 55 | 1707 | 2331 | 0 | 40 | FOEN | 0.5 | − 1.16 | 0.62 | 1.7 | 1.91 | − 0.1 | 0.13 | 0.07 | 0.26 | 2.04 |

**Table 5** (continued)

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner | $I_{BF}$ | PDF location | PDF scale | GPD location | GPD scale | GPD shape | GEV location | GEV scale | GEV shape | $\theta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 101 | Petite Glâne–Villars-le-Grand | 85 | 433 | 560 | 0 | 20 | VD | 0.52 | − 0.92 | 0.56 | 4.8 | 7.72 | − 0.36 | 0.44 | 0.18 | 0.17 | 2.24 |
| 102 | Pfaffnern–Vordemwald | 39 | 417 | 517 | 0 | 34 | AG | 0.55 | − 0.47 | 0.32 | 3.2 | 7.22 | − 0.11 | 0.22 | 0.1 | 0.3 | 1.74 |
| 103 | Plessur–Chur | 263 | 573 | 1850 | 0 | 40 | FOEN | 0.57 | − 0.92 | 0.72 | 12 | 12 | − 0.07 | 1.26 | 0.56 | 0.1 | 2.47 |
| 104 | Poschiavino–La Rösa | 14 | 1860 | 2283 | 0 | 40 | FOEN | 0.53 | − 1.12 | 0.56 | 1.4 | 1.94 | − 0.02 | 0.13 | 0.06 | 0.32 | 2.24 |
| 105 | Promenthouse–Gland | 100 | 394 | 1037 | 0 | 28 | FOEN | 0.26 | − 0.68 | 0.79 | 4.8 | 4.62 | 0.24 | 0.65 | 0.22 | 0.32 | 2.79 |
| 106 | Reppisch–Birmensdorf | 24 | 466 | 665 | 0 | 44 | ZH | 0.31 | − 1.15 | 0.73 | 1.6 | 2.24 | 0.18 | 0.2 | 0.08 | 0.32 | 1.85 |
| 107 | Reppisch–Dietikon | 69 | 380 | 594 | 0 | 28 | ZH | 0.45 | − 1.09 | 0.6 | 5.4 | 8.81 | 0.09 | 0.55 | 0.25 | 0.27 | 2.45 |
| 108 | Riale di Pincascia–Lavertezzo | 44 | 536 | 1708 | 0 | 22 | FOEN | 0.34 | − 0.93 | 0.41 | 25.5 | 28.99 | 0.03 | 2.68 | 0.96 | − 0.02 | 1.52 |
| 109 | Rot–Roggwil | 54 | 436 | 586 | 0 | 25 | FOEN | 0.57 | − 0.64 | 0.41 | 4 | 5.88 | − 0.1 | 0.25 | 0.09 | 0.23 | 1.58 |
| 110 | Ruederchen–Schöftland | 19 | 463 | 614 | 0 | 34 | AG | 0.51 | − 0.78 | 0.47 | 1.2 | 2.59 | 0.03 | 0.09 | 0.04 | 0.27 | 1.8 |
| 111 | Scheulte–Vicques | 73 | 463 | 785 | 0 | 22 | FOEN | 0.46 | − 1.01 | 0.48 | 8 | 10.34 | 0.07 | 0.8 | 0.31 | 0.27 | 2.63 |
| 112 | Schmittenbach–Remigen | 13 | 385 | 523 | 0 | 32 | AG | 0.09 | − 0.88 | 0.71 | 0.9 | 0.98 | 0.12 | 0.15 | 0.06 | 0.23 | 2.99 |
| 113 | Schwarzenbach–Rickenbach | 15 | 410 | 454 | 0 | 22 | ZH | 0.52 | − 0.72 | 0.4 | 1 | 1.52 | − 0.16 | 0.08 | 0.03 | 0.39 | 2.04 |
| 114 | Sellenbodenbach–Neuenkirch | 11 | 515 | 615 | 0 | 23 | FOEN | 0.32 | − 1.14 | 0.53 | 1.5 | 5.16 | 0 | 0.14 | 0.06 | 0.36 | 1.84 |
| 115 | Sense–Thörishaus | 352 | 555 | 1068 | 0 | 36 | FOEN | 0.48 | − 1.03 | 0.58 | 46 | 58.77 | − 0.12 | 3.4 | 1.11 | 0.14 | 1.25 |
| 116 | Seymaz–Thônex | 37 | 393 | 451 | 0 | 20 | GE | 0.43 | − 0.99 | 0.53 | 3 | 2.92 | − 0.08 | 0.31 | 0.11 | 0.18 | 2.41 |
| 117 | Seyon–Valangin | 112 | 630 | 970 | 0 | 34 | FOEN | 0.36 | − 1.01 | 0.49 | 5.8 | 8.85 | − 0.31 | 0.62 | 0.22 | 0.22 | 2.22 |
| 118 | Simme–Oberried/Lenk | 36 | 1096 | 2370 | 35 | 40 | FOEN | 0.34 | − 0.99 | 0.71 | 5.2 | 4.62 | − 0.12 | 0.59 | 0.28 | − 0.08 | 1.71 |
| 119 | Simme–Oberwil | 344 | 777 | 1640 | 4 | 40 | FOEN | 0.58 | − 1.15 | 0.55 | 30.7 | 18.24 | − 0.02 | 2.33 | 0.83 | 0.11 | 2 |
| 120 | Simme–Zweisimmen | 203 | 930 | 1801 | 6 | 21 | BE | 0.61 | − 1.06 | 0.65 | 20.2 | 12.99 | − 0.07 | 1.46 | 0.58 | 0.21 | 2.03 |
| 121 | Sinserbach–Sins | 16 | 415 | 561 | 0 | 33 | AG | 0.33 | − 1.45 | 0.71 | 1.3 | 3.01 | 0.12 | 0.14 | 0.06 | 0.22 | 2.01 |
| 122 | Sionge–Vuippens | 45 | 681 | 862 | 0 | 39 | FOEN | 0.4 | − 1.11 | 0.62 | 7 | 10.98 | − 0.24 | 0.68 | 0.23 | 0.09 | 1.58 |
| 123 | Sissle–Eiken | 123 | 314 | 529 | 0 | 37 | AG | 0.19 | − 0.98 | 0.62 | 9.2 | 11.68 | − 0.05 | 1.16 | 0.43 | 0.17 | 1.85 |
| 124 | Sissle–Hornussen | 37 | 365 | 524 | 0 | 35 | AG | 0.26 | − 1.06 | 0.62 | 3 | 4.82 | − 0.1 | 0.38 | 0.14 | 0.2 | 1.59 |
| 125 | Somvixer Rhein–Somvix | 22 | 1490 | 2450 | 7 | 36 | FOEN | 0.53 | − 0.69 | 0.46 | 6.8 | 5.94 | 0.18 | 0.49 | 0.2 | 0.02 | 1.5 |
| 126 | Sorne–Delémont | 241 | 406 | 808 | 0 | 31 | FOEN | 0.44 | − 0.96 | 0.7 | 17.1 | 7.75 | 0.03 | 1.83 | 0.51 | 0.17 | 2.38 |
| 127 | Staffeleggbach–Frick | 21 | 358 | 534 | 0 | 35 | AG | 0.3 | − 1.01 | 0.69 | 1.5 | 1.96 | − 0.02 | 0.18 | 0.07 | 0.15 | 1.69 |
| 128 | Steinach–Steinach | 24 | 406 | 710 | 0 | 30 | FOEN | 0.53 | − 1.06 | 0.85 | 4.4 | 8.3 | − 0.09 | 0.28 | 0.1 | 0.25 | 1.46 |
| 129 | Steinenbach–Kaltbrunn | 19 | 451 | 1112 | 0 | 30 | FOEN | 0.37 | − 0.74 | 0.45 | 6.4 | 15.9 | − 0.54 | 0.56 | 0.19 | 0.09 | 1.54 |
| 130 | Stichbach–Bottighofen | 16 | 410 | 522 | 0 | 22 | TG | 0.31 | − 1.39 | 0.6 | 1.1 | 2.03 | 0.3 | 0.13 | 0.06 | 0.41 | 3.06 |
| 131 | Surb–Döttingen | 67 | 335 | 511 | 0 | 34 | AG | 0.58 | − 0.74 | 0.42 | 3.8 | 4.88 | 0.05 | 0.26 | 0.11 | 0.28 | 2.01 |
| 132 | Surb–Unterehrendingen | 28 | 424 | 541 | 0 | 21 | AG | 0.59 | − 0.3 | 0.37 | 2.2 | 3.38 | 0.07 | 0.15 | 0.06 | 0.28 | 2.2 |

**Table 5** (continued)

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner | $I_{BF}$ | PDF location | PDF scale | GPD location | GPD scale | GPD shape | GEV location | GEV scale | GEV shape | $\theta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 133 | Suze–Sonceboz | 150 | 642 | 1050 | 0 | 53 | FOEN | 0.43 | − 0.95 | 0.59 | 15.4 | 4.97 | 0.14 | 1.53 | 0.47 | 0.11 | 2.87 |
| 134 | Tägerbach-Wislikofen | 14 | 390 | 551 | 0 | 32 | AG | 0.71 | − 0.9 | 0.79 | 0.4 | 0.46 | − 0.21 | 0.02 | 0.01 | 0.2 | 1.41 |
| 135 | Talbach–Schinznach-Dorf | 15 | 360 | 552 | 0 | 34 | AG | 0.19 | − 0.97 | 0.76 | 0.6 | 1.4 | 0.22 | 0.07 | 0.03 | 0.05 | 1.5 |
| 136 | Talent–Chavornay | 66 | 440 | 670 | 0 | 20 | VD | 0.43 | − 1.05 | 0.54 | 8 | 10.81 | − 0.28 | 0.76 | 0.26 | 0.12 | 1.67 |
| 137 | Taschinasbach-Grüsch | 63 | 666 | 1768 | 0 | 34 | FOEN | 0.39 | − 1.23 | 0.67 | 9.2 | 9.63 | 0.07 | 0.92 | 0.39 | 0.09 | 2.05 |
| 138 | Thur–Andelfingen | 1696 | 356 | 770 | 0 | 40 | FOEN | 0.49 | − 0.85 | 0.44 | 230.9 | 142.25 | − 0.1 | 19.05 | 5.77 | 0.25 | 2.21 |
| 139 | Thur–Halden | 1085 | 456 | 910 | 0 | 40 | FOEN | 0.49 | − 1.16 | 0.58 | 205.2 | 178.62 | − 0.13 | 16.28 | 5.52 | 0.17 | 1.93 |
| 140 | Thur–Jonschwil | 493 | 534 | 1030 | 0 | 40 | FOEN | 0.42 | − 0.87 | 0.42 | 113.1 | 116.86 | − 0.2 | 10.36 | 3.17 | 0.18 | 1.96 |
| 141 | Thur–Stein | 84 | 850 | 1448 | 0 | 31 | FOEN | 0.36 | − 0.92 | 0.46 | 17.2 | 14.14 | 0.02 | 1.93 | 0.62 | 0.23 | 2.6 |
| 142 | Töss–Altlandenberg | 67 | 621 | 871 | 0 | 36 | ZH | 0.24 | − 1.03 | 0.61 | 12.6 | 14.8 | 0.01 | 1.51 | 0.53 | 0.16 | 2.19 |
| 143 | Töss–Freienstein | 399 | 360 | 626 | 0 | 29 | ZH | 0.51 | − 1.13 | 0.6 | 36.4 | 26.99 | 0.14 | 3.54 | 1.19 | 0.17 | 2.72 |
| 144 | Töss–Neftenbach | 342 | 389 | 650 | 0 | 40 | FOEN | 0.48 | − 0.98 | 0.52 | 34.8 | 30.17 | 0.04 | 3.53 | 1.19 | 0.17 | 2.65 |
| 145 | Töss–Rämismühle | 127 | 524 | 790 | 0 | 35 | ZH | 0.31 | − 1.04 | 0.56 | 20.4 | 19.06 | − 0.05 | 2.37 | 0.73 | 0.18 | 2.07 |
| 146 | Trübbach–Räzliberg | 20 | 1430 | 2610 | 54 | 22 | FOEN | 0.2 | − 1.38 | 0.68 | 2.9 | 2.89 | − 0.19 | 0.28 | 0.14 | 0.05 | 1.13 |
| 147 | Ürke–Holziken | 25 | 438 | 577 | 0 | 35 | AG | 0.66 | − 0.86 | 0.53 | 1.3 | 1.97 | − 0.23 | 0.07 | 0.03 | 0.27 | 1.36 |
| 148 | Urnäsch–Hundwil | 65 | 746 | 1085 | 0 | 33 | FOEN | 0.4 | − 1.02 | 0.51 | 16.8 | 20.69 | − 0.05 | 1.46 | 0.51 | 0.09 | 1.49 |
| 149 | Vedeggio–Bioggio | 95 | 280 | 950 | 0 | 35 | FOEN | 0.33 | − 0.89 | 0.64 | 15.2 | 24.87 | 0.11 | 2 | 0.79 | 0.25 | 1.67 |
| 150 | Venoge–Ecublens | 231 | 383 | 700 | 0 | 35 | FOEN | 0.4 | − 0.9 | 0.62 | 16 | 14.37 | − 0.07 | 2.15 | 0.69 | 0.28 | 3.51 |
| 151 | Verzasca–Lavertezzo | 186 | 490 | 1672 | 0 | 24 | FOEN | 0.45 | − 0.84 | 0.4 | 74.3 | 124.12 | − 0.08 | 7.55 | 2.7 | − 0.12 | 1.8 |
| 152 | Veveyse–Vevey | 62 | 425 | 1108 | 0 | 30 | FOEN | 0.39 | − 1.08 | 0.64 | 12.4 | 21.31 | − 0.13 | 1.01 | 0.37 | 0 | 1.27 |
| 153 | Violenbach–Augst | 17 | 268 | 425 | 0 | 35 | BL | 0.29 | − 1.08 | 0.73 | 0.9 | 1.39 | − 0.02 | 0.1 | 0.04 | 0.14 | 1.41 |
| 154 | Vordere Frenke–Bubendorf | 46 | 371 | 647 | 0 | 36 | BL | 0.46 | − 1.02 | 0.5 | 3.6 | 4.44 | − 0.12 | 0.33 | 0.12 | 0.18 | 1.67 |
| 155 | Vordere Frenke–Waldenburg | 13 | 524 | 826 | 0 | 35 | BL | 0.24 | − 0.82 | 0.58 | 1.1 | 1.53 | − 0.17 | 0.12 | 0.05 | 0.25 | 1.42 |
| 156 | Weisse Lütschine–Zweilütschinen | 164 | 650 | 2170 | 18 | 40 | FOEN | 0.56 | − 1.09 | 0.77 | 21.6 | 15.31 | − 0.15 | 1.75 | 0.65 | − 0.09 | 1.48 |
| 157 | Werrikerbach–Greifensee | 12 | 440 | 493 | 0 | 21 | ZH | 0.49 | − 0.99 | 1.17 | 0.4 | 0.18 | − 0.01 | 0.03 | 0.01 | − 0.1 | 1.42 |
| 158 | Wigger–Zofingen | 368 | 426 | 660 | 0 | 35 | FOEN | 0.56 | − 0.85 | 0.53 | 21.4 | 23.91 | − 0.11 | 1.52 | 0.6 | 0.23 | 1.68 |
| 159 | Wissenbach–Boswil | 12 | 460 | 684 | 0 | 34 | AG | 0.39 | − 1.17 | 0.61 | 0.7 | 1 | 0.2 | 0.07 | 0.03 | 0.2 | 1.61 |
| 160 | Wölflinswiler Bach–Wittnau | 17 | 395 | 600 | 0 | 30 | AG | 0.39 | − 1 | 0.59 | 1.5 | 3.03 | − 0.19 | 0.16 | 0.06 | 0.15 | 1.69 |
| 161 | Wyna–Reinach | 47 | 514 | 682 | 0 | 30 | AG | 0.53 | − 0.39 | 0.32 | 4 | 4.86 | − 0.02 | 0.31 | 0.12 | 0.22 | 1.86 |
| 162 | Wyna–Suhr | 120 | 392 | 617 | 0 | 34 | AG | 0.5 | − 0.91 | 0.61 | 6.9 | 9.68 | − 0.21 | 0.6 | 0.24 | 0.28 | 1.82 |

**Table 5** (continued)

| ID | Station name | Area | ELEV | MELEV | DG | RL | Owner | $I_{BF}$ | PDF location | PDF scale | GPD location | GPD scale | GPD shape | GEV location | GEV scale | GEV shape | $\theta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 163 | Wyna–Unterkulm | 92 | 455 | 649 | 0 | 37 | AG | 0.53 | − 0.79 | 0.48 | 7 | 9.29 | − 0.18 | 0.56 | 0.21 | 0.26 | 1.62 |

The station name is provided together with a catchment ID, catchment area [$km^2$], elevation [m.a.s.l.] (ELEV), mean elevation [m.a.s.l.] (MELEV), degree of glaciation [%] (DG), record length [a] (RL), and the owner of the station. Data owners are the Federal Office for the Environment (FOEN), and different cantons: Zürich (ZH), Vaud (VD), Solothurn (SO), Bern (BE), Baselland (BL), Aargau (AG), and Thurgau (TG). The SDH parameters are explained in Fig. 2

## References


Abrahart RJ, See LM (2007) Neural network modelling of non-linear hydrological relationships. Hydrol Earth Syst Sci 11:1563–1579. https://doi.org/10.5194/hess-11-1563-2007

Acreman MC, Sinclair CD (1986) Classification of drainage basins acording to their physical characteristics; an application for flood frequency analysis in Scotland. J Hydrol 84:365–380. https://doi.org/10.1016/0022-1694(86)90134-4

Ahn KH, Palmer R (2016) Regional flood frequency analysis using spatial proximity and basin characteristics: quantile regression versus parameter regression technique. J Hydrol 540:515–526. https://doi.org/10.1016/j.jhydrol.2016.06.047

Ali G, Tetzlaff D, Soulsby C, McDonnell JJ, Capell R (2012) A comparison of similarity indices for catchment classification using a cross-regional dataset. Adv Water Resour 40:11–22. https://doi.org/10.1016/j.advwatres.2012.01.008

Archfield SA, Pugliese A, Castellarin A, Skøien JO, Kiang JE (2013) Topological and canonical kriging for design flood prediction in ungauged catchments: an improvement over a traditional regional regression approach? Hydrol Earth Syst Sci 17(4):1575–1588. https://doi.org/10.5194/hess-17-1575-2013

Aziz K, Rai S, Rahmen A (2015) Design flood estimation in ungauged catchments using genetic algorithm-based artificial neural network (GAANN) technique for Australia. Nat Hazards 77:805–821. https://doi.org/10.1007/s11069-015-1625-x

Aziz K, Haque MM, Rahman A, Shamseldin AY, Shoaib M (2016) Flood estimation in ungauged catchments: application of artificial intelligence based methods for Eastern Australia. Stoch Env Res Risk Assess. https://doi.org/10.1007/s00477-016-1272-0

Bardossy A (2007) Calibration of hydrological model parameters for ungauged catchments. Hydrol Earth Syst Sci 11:703–710

Bardossy A, Lehmann W (1997) Spatial distribution of soil moisture in a small catchment. Part 1: geostatistical analysis. J Hydrol 206:1–15. https://doi.org/10.1016/S0022-1694(97)00152-2

Beygelzimer A, Kakadet S, Langford J, Arya S, Mount D, Li S (2013) Package 'FNN': fast nearest neighbor search algorithms and applications. http://cran.r-project.org/package=FNN

Bhunya PK, Panda SN, Goel MK (2011) Synthetic unit hydrograph methods: a critical review. Open Hydrol J 5:1–8. https://doi.org/10.2174/1874378101105010001

Bitterli T, Aviolat P, Brändli R, Christe R, Fracheboud S, Frey D, George M, Matousek F, Tripet JP (2007) Groundwater resources. In: Hydrological Atlas of Switzerland, Bern, p 8.6

Blöschl G (2006) Geostatistische Methoden bei der hydrologischen Regionalisierung. In: Godina R, Blöschl G (eds) Methoden der hydrologischen Regionalisierung, vol 197. Wiener Mitteilungen, Wien, pp 21–40

Blöschl G, Sivapalan M, Wagener T, Viglione A, Savenije H (2013) Runoff prediction in ungauged basins. Cambridge University Press, Cambridge

Boscarello L, Ravazzani G, Cislaghi A, Mancini M (2016) Regionalization of flow-duration curves through catchment classification with streamflow signatures and physiographic-climate indices. J Hydrol Eng. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001307

Breiman L (1996) Bagging predictors. Mach Learn 24(421):123–140. https://doi.org/10.1007/BF00058655

Breiman L (2001) Random Forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Brunner MI, Seibert J, Favre AC (2016) Bivariate return periods and their importance for flood peak and volume estimation. Wire's Water 3:819–833. https://doi.org/10.1002/wat2.1173

Brunner MI, Sikorska AE, Furrer R, Favre AC (2017a) Uncertainty assessment of synthetic design hydrographs for gauged and ungauged catchments. Water Resour Res **(accepted)**

Brunner MI, Viviroli D, Sikorska AE, Vannier O, Favre AC, Seibert J (2017b) Flood type specific construction of synthetic design hydrographs. Water Resour Res. https://doi.org/10.1002/2016WR019535

Bundesamt für Statistik (2003) Geodaten der Bundesstatistik. https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik.html

Burn DH (1989) Cluster analysis as applied to regional flood frequency. J Water Resour Plan Manag 115:567–582

Burn DH (1990) Evaluation of regional flood frequency analysis with a region of influence approach. Water Resour Res 26(10):2257–2265. https://doi.org/10.1029/WR026i010p02257

Burn DH, Boorman DB (1992) Catchment classification applied to the estimation of hydrological parameters at ungauged catchments. Tech. rep, Institute of Hydrology, Wallingford, Oxfordshire

Burn DH, Boorman DB (1993) Estimation of hydrological parameters at ungauged catchments. J Hydrol 143:429–454. https://doi.org/10.1016/0022-1694(93)90203-L

Camezind-Wildi R (2005) Empfehlung Raumplanung und Naturgefahren. Tech. rep., Bundesamt für Raumentwicklung, Bundesamt für Wasser und Geologie, Bundesamt für Umwelt, Wald und Landschaft, Bern

Castellarin A, Burn DH, Brath A (2001) Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. J Hydrol 241(3):270–285. https://doi.org/10.1016/S0022-1694(00)00383-8

Castiglioni S, Castellarin A, Montanari A, Skøien JO, Laaha G, Blöschl G (2011) Smooth regional estimation of low-flow indices: physiographical space based interpolation and top-kriging. Hydrol Earth Syst Sci 15(3):715–727. https://doi.org/10.5194/hess-15-715-2011

Cavadias GS, Ouarda TBMJ, Bobée B, Girard C (2001) A canonical correlation approach to the determination of homogeneous regions for regional flood estimation of ungauged basins. Hydrol Sci J 46(4):499–512. https://doi.org/10.1080/02626660109492846

Centre for Ecology and Hydrology (1999) Flood estimation handbook. Centre for Ecology and Hydrology, Wallingford, Oxfordshire

Chapman TG, Maxwell AI (1996) Baseflow separation–comparison of numerical methods with tracer experiments. 23rd hydrology and water resources symposium. Hobart, Australia, pp 539–545

Chebana F, Ouarda T (2009) Multivariate quantiles in hydrological frequency analysis. Environmetrics 22:63–78. https://doi.org/10.1002/env.1027

Cheng L, Yaeger M, Viglione A, Coopersmith E, Ye S, Sivapalan M (2012) Exploring the physical controls of regional patterns of flow duration curves—Part 1: insights from statistical analyses. Hydrol Earth Syst Sci 16:4435–4446. https://doi.org/10.5194/hess-16-4435-2012

Chokmani K, Ouarda TBMJ (2004) Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. Water Resour Res 40(12):W12,514. https://doi.org/10.1029/2003WR002983

Cipriani T, Toilliez T, Sauquet E (2012) Estimation régionale des débits décennaux et durées caractéristiques de crue en France. La Houille Blanche 4–5:5–13. https://doi.org/10.1051/lhb/2012024

Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London

Cuevas A, Febrero M, Fraiman R (2007) Robust estimation and classification for functional data via projection-based depth notions. Comput Stat 22(3):481–496. https://doi.org/10.1007/s00180-007-0053-0

Dawson CW, Abrahart RJ, Shamseldin AY, Wilby RL (2006) Flood estimation at ungauged sites using artificial neural networks. J Hydrol 319:391–409. https://doi.org/10.1016/j.jhydrol.2005.07.032

Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall (2012) Merkblatt DWA-M 552. Tech. rep, DWA, Hennef, Germany

Diggle PJ, Ribeiro PJ Jr (2007) Model-based geostatistics. Springer series in statistics. Springer, New York

Eckhardt K (2005) How to construct recursive digital filters for baseflow separation. Hydrol Process 19:507–515. https://doi.org/10.1002/hyp.5675

Eidgenössische Forschungsanstalt für Wald Schnee und Landschaft (WSL) (1999) Schweizerisches Landesforstinventar. Ergebnisse der Zwietaufnahme 1993-1995. BUWAL, Bern

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77:802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Farmer WH (2016) Ordinary kriging as a tool to estimate historical daily streamflow records. Hydrol Earth Syst Sci 20(7):2721–2735. https://doi.org/10.5194/hess-20-2721-2016

Freund Y, Schapire RRE (1996) Experiments with a new boosting algorithm. In: International conference on machine learning, pp 148–156, https://doi.org/10.1.1.133.1040, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.6252

Friedman AJ, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordiante descent. J Stat Softw 33(1):1–22

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232

Friedman JH (2002) Stochastic gradient boosting. Comput Stat Data Anal 38(4):367–378

Gaál L, Kysel J, Szolgay J (2008) Region-of-influence approach to a frequency analysis of heavy precipitation in Slovakia. Hydrol Earth Syst Sci 12:825–839. https://doi.org/10.5194/hess-12-825-2008

Ganora D, Claps P, Laio F, Viglione A (2009) An approach to estimate nonparametric flow duration curves in ungauged basins. Water Resour Res 45(10):1–10. https://doi.org/10.1029/2008WR007472

Genest C, Favre AC (2007) Everything you always wanted to know about copula modeling but were afraid to ask. J Hydrol Eng 12(4):347–368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

Gottschalk L (1993) Correlation and covariance of runoff. Stoch Hydrol Hydraul 7:85–101

Gottschalk L, Leblois E, Skoien JO (2011) Correlation and covariance of runoff revisited. J Hydrol 398:76–90. https://doi.org/10.1016/j.jhydrol.2010.12.011

Green IRA, Stephenson D (1986) Criteria for comparison of single event models. Hydrol Sci J 31(3):395–411. https://doi.org/10.1080/02626668609491056

Greene W (2002) Econometric analysis, 5th edn. Prentice Hall, New Jersey

GREHYS (1996) Presentation and review of some methods for regional flood frequency analysis. J Hydrol 186:63–84. https://doi.org/10.1016/S0022-1694(96)03042-9

Grimaldi S, Petroselli A (2015) Do we still need the Rational Formula? An alternative empirical procedure for peak discharge estimation in small and ungauged basins. Hydrol Sci J 60(1):67–77. https://doi.org/10.1080/02626667.2014.880546

Haberlandt U, Klöcking B, Krysanova V, Becker A (2001) Regionalisation of the base flow index from dynamically simulated flow

components—a case study in the Elbe River Basin. J Hydrol 248:35–53. https://doi.org/10.1016/S0022-1694(01)00391-2

Haddad K, Rahman A (2012) Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework—quantile regression vs. parameter regression technique. J Hydrol 430–431:142–161. https://doi.org/10.1016/j.jhydrol.2012.02.012

Halkidi M, Batistakis Y, Vazirgiannis M (2001) On clustering validation techniques. J Intell Inf Syst 17(2/3):107–145. https://doi.org/10.1023/A:1012801612483

Harrell FE (2015) Regression modeling strategies. With applications to linear models, logistic and ordinal regression, and survival analysis. Springer, Cham

Hastie T, Tibshirani R, Friedman J (2008) The elements of statistical learning. Springer series in statistics. Springer, Stanford

He Y, Bardossy A, Zehe E (2011) A review of regionalisation for continuous streamflow simulation. Hydrol Earth Syst Sci 15:3539–3553. https://doi.org/10.5194/hess-15-3539-2011

Hechenbichler K, Schliep K (2004) Weighted k-nearest-neighbor techniques and ordinal classification. Mol Ecol 399:17

Hofner B, Mayr A, Robinzonov N, Schmid M (2009) Model-based Boosting in R. A Hands-on Tutorial Using the R Package mboost. Tech. rep., Department of statistics. University of Munich, Munich

Hosking JRM, Wallis JR (1993) Some statistics useful in regional frequency analysis. Water Resour Res 29(92):271–281

Hrachowitz M, Savenije HHG, Blöschl G, McDonnell JJ, Sivapalan M, Pomeroy JW, Arheimer B, Blume T, Clark MP, Ehret U, Fenicia F, Freer JE, Gelfan A, Gupta HV, Hughes DA, Hut RW, Montanari A, Pande S, Tetzlaff D, Troch PA, Uhlenbrook S, Wagener T, Winsemius HC, Woods RA, Zehe E, Cudennec C (2013) A decade of Predictions in Ungauged Basins (PUB)—a review. Hydrol Sci J 58(6):1198–1255. https://doi.org/10.1080/02626667.2013.803183

Hundecha Y, Ouarda TBMJ, Bardossy A (2008) Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the spatial structures of the parameters within a canonical physiographic-climatic space. Water Resour Res 44(W01):427. https://doi.org/10.1029/2006WR005439

Ilorme F, Griffis VW (2013) A novel procedure for delineation of hydrologically homogeneous regions and the classification of ungauged sites for design flood estimation. J Hydrol 492:151–162. https://doi.org/10.1016/j.jhydrol.2013.03.045

James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning. With applications in R. Springer, New York. https://doi.org/10.1007/978-1-4614-7138-7

Jensen H, Lang H, Rinderknecht J (1997) Extreme point rainfall of varying duration and return period 1901–1970. In: Hydrological Atlas of Switzerland, FOEN, Bern, chap 2.4

Ji Z, Li N, Xie W, Wu J, Zhou Y (2013) Comprehensive assessment of flood risk using the classification and regression tree method. Stoch Env Res Risk Assess 27(8):1815–1828. https://doi.org/10.1007/s00477-013-0716-z

Joe H (1997) Multivariate models and dependence concepts. Chapman and Hall/CRC, London

Kiers HAL, Smilde AK (2007) A comparison of various methods for multivariate regression with highly collinear variables. Stat Methods Appl 16:193–228. https://doi.org/10.1007/s10260-006-0025-5

Kjeldsen TR, Jones DA (2010) Predicting the index flood in ungauged UK catchments: on the link between data-transfer and spatial model error structure. J Hydrol 387(1–2):1–9. https://doi.org/10.1016/j.jhydrol.2010.03.024

Kokkonen TS, Jakeman AJ, Young PC, Koivusalo HJ (2003) Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. Hydrol Process 17(11):2219–2238. https://doi.org/10.1002/hyp.1329

Laaha G, Blöschl G (2006) A comparison of low flow regionalisation methods-catchment grouping. J Hydrol 323(1–4):193–214. https://doi.org/10.1016/j.jhydrol.2005.09.001

Laaha G, Skoien JO, Blöschl G (2014) Spatial prediction on river networks: comparison of top-kriging with regional regression. Hydrol Process 28:315–324. https://doi.org/10.1002/hyp.9578

Lang M, Ouarda T, Bobée B (1999) Towards operational guidelines for over-threshold modeling. J Hydrol 225:103–117

Le Cessie S, van Houwelingen JC (1992) Ridge estimators in logistic regression. J Appl Stat 41(1):191–201. https://doi.org/10.2307/2347628

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3):18–22

Longobardi A, Villani P (2008) Baseflow index regionalization analysis in a Mediterranean area and data scarcity context: role of the catchment permeability index. J Hydrol 355:63–75. https://doi.org/10.1016/j.jhydrol.2008.03.011

Lu GY, Wong DW (2008) An adaptive inverse-distance weighting spatial interpolation technique. Comput Geosci 34(9):1044–1055. https://doi.org/10.1016/j.cageo.2007.07.010

Matheron G (1971) The theory of regionalized variables and its applications, vol 5. École nationale supérieure des Mines, Paris

McIntyre N, Lee H, Wheater H, Young A, Wagener T (2005) Ensemble predictions of runoff in ungauged catchments. Water Resour Res 41(12):1–14. https://doi.org/10.1029/2005WR004289

Mediero L, Jiménez-Alvarez A, Garrote L (2010) Design flood hydrographs from the relationship between flood peak and volume. Hydrol Earth Syst Sci 14:2495–2505. https://doi.org/10.5194/hess-14-2495-2010

Merz R (2006) Regionalisierung von statistischen Hochwasserkenngrössen. In: Godina R, Blöschl G (eds) Methoden der hydrologischen Regionalisierung, vol 197. Wiener Mitteilungen, Wien, pp 109–130

Merz R, Blöschl G (2003) A process typology of regional floods. Water Resour Res 39(12):1340. https://doi.org/10.1029/2002WR001952

Merz R, Blöschl G (2004) Regionalisation of catchment model parameters. J Hydrol 287(1):95–123. https://doi.org/10.1016/j.jhydrol.2003.09.028

MeteoSwiss (2013) Documentation of MeteoSwiss grid-data products: Daily precipitation (final analysis): RhiresD. Tech. rep., MeteoSwiss, http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/precip/rhiresd.Par.0007.DownloadFile.tmp/proddocrhiresd.pdf

Mevik BH, Wehrens R (2007) The pls package: principal component and partial least squares regression in R. J Stat Softw 18(2):1–23. https://doi.org/10.18637/jss.v018.i02

Meylan P, Favre AC, Musy A (2012) Predictive hydrology. A frequency analysis approach. Science Publishers, St. Helier, Jersey, British Channel Islands

Myers RH, Montgomery DC, Vining GG, Robinson TJ (2010) Generalized Linear Models, vol 4. Wiley, Hoboken

Nathan RJ, McMahon TA (1990) Identification of homogeneous regions for the purposes of regionalisation. J Hydrol 121:217–238

Nied M, Pardowitz T, Nissen K, Ulbrich U, Hundecha Y, Merz B (2014) On the relationship between hydro-meteorological patterns and flood types. J Hydrol 519:3249–3262. https://doi.org/10.1016/j.jhydrol.2014.09.089

Osborne JW (2010) Improving your data transformations: applying the Box-Cox transformation. Pract Assess Res Eval 15(12):1–9

Ouarda T, Cunderlik JM, St-Hilaire A, Barbet M, Bruneau P, Bobée B (2006) Data-based comparison of seasonality-based regional

flood frequency methods. J Hydrol 330(1):329–339. https://doi.org/10.1016/j.jhydrol.2006.03.023
Ouarda TBMJ, Haché M, Bruneau P, Bobée B (2000) Regional flood peak and volume estimation in northern Canadian basin. J Cold Reg Eng 14:176–191. https://doi.org/10.1061/(ASCE)0887-381X(2000)14:4(176)
Ouarda TBMJ, Girard C, Cavadias GS, Bobée B (2001) Regional flood frequency estimation with canonical correlation analysis. J Hydrol 254:157–173
Oudin L, Andréassian V, Perrin C, Michel C, Moine NL (2008) Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. Water Resour Res 44:W03, 413. https://doi.org/10.1029/2007WR006240
Oudin L, Kay A, Andréassian V, Perrin C (2010) Are seemingly physically similar catchments truly hydrologically similar? Water Resour Res 46(W11):558. https://doi.org/10.1029/2009WR008887
Parajka J, Merz R, Blöschl G (2005) A comparison of regionalisation methods for catchment model parameters. Hydrol Earth Syst Sci 9:157–171. https://doi.org/10.5194/hess-9-157-2005
Parajka J, Andréassian V, Archfield SA, Bardossy A, Blöschl G, Chiew F, Duan Q, Gelfan A, Hlavconva K, Merz R, McIntyre N, Oudin L, Perrin C, Rogger M, Salinas JL, Savenije HG, Skoien JO, Wagener T, Zehe E, Zhang Y (2013) Prediction of runoff hydrographs in ungauged basins. In: Blöschl G, Sivapalan M, Wagener T, Viglione A, Savenije H (eds) Predictions in ungauged basins. A synthesis across processes, places and scales, Cambridge University Press, Cambridge, pp 227–269
Pebesma E (2004) Multivariable geostatistics in S: the gstat package. Comput Geosci 30:683–691. https://doi.org/10.1016/j.cageo.2004.03.012
Peters A, Hothorn T, Ripley BD, Therneau T, Atkinson B (2015) Package ' ipred ' : improved predictors. http://cran.r-project.org/package=ipred
Petroselli A, Grimaldi S (2015) Design hydrograph estimation in small and fully ungauged basins: a preliminary assessment of the EBA4SUB framework. J Flood Risk Manag. https://doi.org/10.1111/jfr3.12193
Pilgrim DH (1986) Bridging the gap between flood research and design practice. Water Resour Res 22(9):165–176
Prinzio MD, Castellarin A, Toth E (2011) Data-driven catchment classification: application to the pub problem. Hydrol Earth Syst Sci 15:1921–1935. https://doi.org/10.5194/hess-15-1921-2011
R Core Team (2015) R: a language and environment for statistical computing. http://www.r-project.org/
Rahman A, Charron C, Ouarda TBMJ, Chebana F (2017) Development of regional flood frequency analysis techniques using generalized additive models for Australia. Stoch Environ Res Risk Assess pp 1–17, https://doi.org/10.1007/s00477-017-1384-1
Rai RK, Sarkar S, Singh VP (2009) Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph. Water Resour Manage 23:899–929. https://doi.org/10.1007/s11269-008-9306-0
Rasmussen PF, Bobée B, Bernier J (1993) Une méthodologie générale de comparaison de modèles d'estimation régionale de crue. Revue des sciences de l'eau 7:23–41. https://doi.org/10.7202/705187ar
Razavi T, Coulibaly P (2013) Streamflow prediction in ungauged basins: review of regionalization methods. J Hydrol Eng 18(8):958–975. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000690
Requena AI, Mediero L, Garrote L (2013) A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation. Hydrol Earth Syst Sci 17:3023–3038. https://doi.org/10.5194/hess-17-3023-2013
Ridgeway G (2007) Generalized boosted models: a guide to the gbm package. Compute 1(4):1–12. https://doi.org/10.1111/j.1467-9752.1996.tb00390.x
Rosbjerg D, Blöschl G, Burn DH, Castellarin A, Croke B, Baldassarre GD, Iacobellis V, Kjeldsen TR, Kuczera G, Merz R, Montanari A, Morris D, Ouarda T, Ren L, Rogger M, Salinas JL, Toth E, Viglione A (2013) Prediction of floods in ungauged basins. In: Blöschl G, Sivapalan M, Wagener T, Viglione A, Savenije H (eds) Runoff prediction in ungauged basins. A synthesis across processes, places and scales, Cambridge University Press, Cambridge, chap 9, pp 189–226
Salinas JL, Laaha G, Rogger M, Parajka J, Viglione A, Sivapalan M, Blöschl G (2013) Comparative assessment of predictions in ungauged basins—Part 2: flood and low flow studies. Hydrol Earth Syst Sci 17:2637–2652. https://doi.org/10.5194/hess-17-2637-2013
Samuel J, Coulibaly P, Metcalfe RA (2011) Estimation of continuous streamflow in ontario ungauged basins: comparison of regionalization methods. J Hydrol Eng 16(5):447–459. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000338
Sauquet E (2006) Mapping mean annual river discharges: geostatistical developments for incorporating river network dependencies. J Hydrol 331:300–314. https://doi.org/10.1016/j.jhydrol.2006.05.018
Sauquet E, Catalogne C (2011) Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France. Hydrol Earth Syst Sci 15:2421–2435. https://doi.org/10.5194/hess-15-2421-2011
Sefton CEM, Howarth SM (1998) Relationships between dynamic response characteristics and physical descriptors of catchments in England and wales. J Hydrol 211(1–4):1–16. https://doi.org/10.1016/S0022-1694(98)00163-2
Serinaldi F, Grimaldi S (2011) Synthetic design hydrographs based on distribution functions with finite support. J Hydrol Eng 16:434–446. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000339
Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(34):591–611
Shiau J, Wang HY, Tsai CT (2006) Bivariate Frequency Analysis of floods using copulas. J Am Water Resour Assoc pp 1549–1564, https://doi.org/10.1111/j.1752-1688.2006.tb06020.x
Shu C, Burn DH (2004) Artificial neural network ensembles and their application in pooled flood frequency analysis. Water Resour Res 40(9):1–10. https://doi.org/10.1029/2003WR002816
Shu C, Ouarda T (2008) Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. J Hydrol 349:31–43. https://doi.org/10.1016/j.jhydrol.2007.10.050
Sikorska AE, Viviroli D, Seibert J (2015) Flood type classification in mountainous catchments using crisp and fuzzy decision trees. Water Resour Res 51(10):7959–7976. https://doi.org/10.1002/2015WR017326
Singh PK, Mishra SK, Jain MK (2014) A review of the synthetic unit hydrograph: from the empirical UH to advanced geomorphological methods. Hydrol Sci J. https://doi.org/10.1080/02626667.2013.870664
Sivapalan M (2003) Prediction in ungauged basins: a grand challenge for theoretical hydrology. Hydrol Process 17:3163–3170. https://doi.org/10.1002/hyp.5155
Skoien JO, Merz R, Blöschl G (2006) Top-kriging—geostatistics on stream networks. Hydrol Earth Syst Sci 10:277–287. https://doi.org/10.5194/hess-10-277-2006
Skoien JO, Blöschl G, Laaha G, Pebesma E, Parajka J, Viglione A (2014) An R package for interpolation of data with a variable spatial support, with an example from river networks. Comput Geosci 67:180–190

Smithers JC (2012) Methods for design flood estimation in South Africa. Water SA 38(4):633–646. https://doi.org/10.4314/wsa.v38i4.19

Steinschneider S, Yang YCE, Brown C (2014) Combining regression and spatial proximity for catchment model regionalization: a comparative study. Hydrol Sci J 6667:1–18. https://doi.org/10.1080/02626667.2014.899701

Strobl C, Malley J, Tutz G (2009) An introduction to recursive partitioning: rationale, application and characteristics of classification and regression trees, bagging and random forests. Psychol Methods 14(4):323–348. https://doi.org/10.1037/a0016973

Takezawa K (2012) Introduction to nonparametric regression. Wiley, Hoboken, https://doi.org/10.1021/cr2001349

Tibshirani R (1997) The lasso method for variable selection in the Cox model. Stat Med 16(4):385–395. https://doi.org/10.1002/(SICI)1097-0258(19970228)16:4%3c385::AID-SIM380%3e3.0.CO;2-3

Tung YK, Yeh KC, Yang JC (1997) Regionalization of unit hydrograph parameters: 1. Comparison of regression analysis techniques. Stoch Hydrol Hydraul 11:145–171

Viglione A, Merz R, Blöschl G (2009) On the role of the runoff coefficient in the mapping of rainfall to flood return periods. Hydrol Earth Syst Sci 6(1):627–665. https://doi.org/10.5194/hessd-6-627-2009

Viviroli D, Mittelbach H, Gurtz J, Weingartner R (2009a) Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland—Part II: parameter regionalisation and flood estimation results. J Hydrol 377:208–225. https://doi.org/10.1016/j.jhydrol.2009.08.022

Viviroli D, Zappa M, Gurtz J, Weingartner R (2009b) An introduction to the hydrological modelling system PREVAH and its pre-and post-processing-tools. Environ Model Softw 24:1209–1222. https://doi.org/10.1016/j.envsoft.2009.04.001

Webster R, Oliver MA (2007) Geostatistics for environmental scientists. Statistics in practice. Wiley, Chichester

Weisberg S (2005) Applied Linear Regression, 3rd edn. Wiley, Hoboken

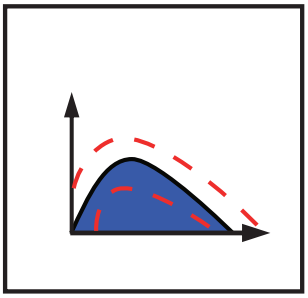
Yamamoto JK (2007) On unbiased backtransform of lognormal kriging estimates. Comput Geosci 11:219–234. https://doi.org/10.1007/s10596-007-9046-x

Yue S, Rasmussen P (2002) Bivariate frequency analysis: discussion of some useful concepts in hydrological application. Hydrol Process 16:2881–2898. https://doi.org/10.1002/hyp.1185

Yue S, Ouarda T, Bobée B, Legendre P, Bruneau P (2002) Approach for describing statistical properties of flood hydrograph. J Hydrol Eng 7(2):147–153. https://doi.org/10.1061/(ASCE)1084-0699(2002)7:2(147)

Zhang Y, Chiew FHS (2009) Relative merits of different methods for runoff predictions in ungauged catchments. Water Resour Res 45(W07):412. https://doi.org/10.1029/2008WR007504

# Paper IV



# Uncertainty assessment of synthetic design hydrographs for gauged and ungauged catchments

Manuela I. Brunner[1,2], Anna E. Sikorska[1,3], Reinhard Furrer[4], and Anne-Catherine Favre[2]

[1]Department of Geography, University of Zurich, Zurich, Switzerland
[2]Univ. Grenoble-Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[3]Department of Hydraulic Engineering, Warsaw University of Life Sciences – SGGW, Warsaw, Poland
[4]Department of Mathematics, University of Zurich, Zurich, Switzerland

*Water Resources Research, accepted*

**Abstract**

Design hydrographs described by peak discharge, hydrograph volume, and hydrograph shape are essential for engineering tasks involving storage. Such design hydrographs are inherently uncertain as are classical flood estimates focusing on peak discharge only. Various sources of uncertainty contribute to the total uncertainty of synthetic design hydrographs for gauged and ungauged catchments. These comprise model uncertainties, sampling uncertainty, and uncertainty due to the choice of a regionalization method. A quantification of the uncertainties associated with flood estimates is essential for reliable decision making and allows for the identification of important uncertainty sources. We therefore propose an uncertainty assessment framework for the quantification of the uncertainty associated with synthetic design hydrographs. The framework is based on bootstrap simulations and consists of three levels of complexity. On the first level, we assess the uncertainty due to individual uncertainty sources. On the second level, we quantify the total uncertainty of design hydrographs for gauged catchments and the total uncertainty of regionalizing them to ungauged catchments but independently from the construction uncertainty. On the third level, we assess the coupled uncertainty of synthetic design hydrographs in ungauged catchments, jointly considering construction and regionalization uncertainty. We find that the most important sources of uncertainty in design hydrograph construction are the record length and the choice of the flood sampling strategy. The total uncertainty of design hydrographs in ungauged catchments depends on the catchment properties and is not negligible in our case.

# 1 Introduction

Hydrograph volume and shape, in addition to peak discharge, are important hydrograph characteristics for flood risk management tasks such as the planning of retention basins and drawing hazard maps (*Tung and Yen*, 2005; *Klein et al.*, 2010; *Schumann et al.*, 2010; *Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012). A complete hydrograph is essential for all designs involving storage (*Pilgrim*, 1986) where the peak discharge, hydrograph volume, and the hydrograph shape provide complementary information. However, flood frequency analyses often focus on peak discharges without considering their dependence on hydrograph volumes. *Brunner et al.* (2017a) therefore proposed an approach to construct synthetic design hydrographs (SDHs) that provide information on the peak discharge and the corresponding hydrograph volume together with the hydrograph shape. This approach takes into account the dependence between peak discharges and hydrograph volumes and models the hydrograph shape via a probability density function (*Yue et al.*, 2002). In a follow up paper, *Brunner et al.* (2017b) assessed how such SDHs can be transferred from gauged to ungauged catchments and identified the most suitable regionalization model. These previous studies suggested, that the construction and regionalization of SDHs may be linked with non-negligible uncertainty which should be quantified in a next step.

Studies involving flood estimation entail various sources of uncertainty such as measurement errors, various assumptions, sample selection, the choice of a suitable distribution function, the choice of a parameter estimation method, and sampling uncertainty (*Merz and Thieken*, 2005). Measurement errors comprise errors in water level measurements and errors coming from transferring water levels into discharge values via a rating curve (*Sikorska et al.*, 2013; *McMillan and Westerberg*, 2015). Assumptions include those of stationarity, homogeneity, and independence of the data. Sample selection is associated with the choice of a representative observation period and the choice of a sampling strategy (annual maxima sampling versus peak-over-threshold sampling). Also, several distribution functions have been used to model the distribution of flood samples and the choice of one suitable distribution over another one is linked to uncertainty. The parameters of such a distribution can be estimated using different estimation techniques such as maximum likelihood and the method of moments or L-moments. Sampling uncertainty describes the uncertainty introduced by not knowing the population underlying a dataset (*Merz and Thieken*, 2005). Among these sources of uncertainty, data availability and model choice are said to be the most important (*Apel et al.*, 2004; *Merz and Thieken*, 2005; *Botto et al.*, 2014). Yet, each step in the modeling process can introduce uncertainty (*Kidson and Richards*, 2005) and the overall uncertainty of the flood estimates results from the interaction of several uncertainty sources (*Merz et al.*, 2008; *Beven and Hall*, 2014), which do not have to be additive (*Montanari and Koutsoyiannis*, 2012). Despite its importance, this uncertainty is often overlooked (*Pappenberger and Beven*, 2006) even though its consideration has several advantages (*Juston et al.*, 2013). Uncertainty analysis allows the identification of uncertain parameters (*Tung and Yen*, 2005), a quantitative assessment of model reliability (*Merz and Thieken*, 2005; *Tung and Yen*, 2005; *Montanari and Koutsoyiannis*, 2012), and it provides a means of analyzing the robustness of flood risk management decisions. Furthermore, an analysis of the contribution of individual sources indicates where potential improvements in the method could have the greatest impact (*Cullen and Frey*, 1999; *Hall and Solomatine*, 2008; *Sikorska et al.*, 2012) and therefore how uncertainty could be reduced (*Qi et al.*, 2016), which is especially important for ungauged catchments (*Sikorska et al.*, 2012). Although, uncertainty can not be eliminated, its assessment at least enables its management (*Koutsoyiannis*, 2014).

Previous studies have dealt mainly with uncertainty analyses for univariate design variable quantiles (*Serinaldi*, 2009) usually estimated based on a sample of peak discharges. There, rather simple analytical and bootstrap methods allow the exploration of the uncertainty of extreme quantiles. In a univariate framework, the effect of the choice of the marginal distribution (*Merz and Thieken*, 2005; *Qi et al.*, 2016), parameter uncertainty of the marginal distribution (*Qi et al.*, 2016), data uncertainty from threshold selection (*Xu et al.*, 2010; *Qi et al.*, 2016), and the effect of the choice of annual maxima sampling versus peak-over-threshold sampling (e.g. *Madsen et al.* (1997); *Martins and Stedinger* (2001a,b); *Sun et al.* (2017)) have been considered. In a bivariate framework that allows for the joint consideration of peak discharges and hydrograph volumes, the effect of the choice of annual maxima sampling versus peak-over-threshold sampling has, to our knowledge, not yet been analyzed. Furthermore, the uncertainty of the marginal distributions combines with the uncertainty of their dependence structure and infinite combinations of the studied variables exist that share the same joint

probability (*Serinaldi*, 2013). Recently, *Serinaldi* (2013) and *Dung et al.* (2015) proposed several parametric and non parametric bootstrap algorithms to compute confidence intervals for bivariate quantiles. However, there is still a lack of understanding of combined and interactive contributions of different uncertainty sources in bivariate quantile estimation (*Qi et al.*, 2016). Furthermore, it is not clear how the uncertainty of bivariate design estimates describing the magnitude of an event interacts with the uncertainty related to the hydrograph shape. The goal of this study is therefore threefold:

1. Assessing the effect of the choice of a peak-over-threshold versus an annual maxima sampling strategy on design hydrograph construction in a bivariate framework.

2. Identifying the most important sources of uncertainty in design hydrograph construction and regionalization.

3. Assessing the uncertainty of synthetic design hydrographs for gauged and ungauged catchments.

To answer these questions, we propose an uncertainty assessment framework based on simulations with three levels of complexity. In a first step, the effect of different uncertainty sources on SDH construction is assessed. This allows for the identification of relevant uncertainty sources and therefore enables the refinement of the SDH construction (*Brunner et al.*, 2017a) and regionalization procedures (*Brunner et al.*, 2017b) in order to reduce uncertainty. Then, we assess the total uncertainty of SDHs for gauged and ungauged catchments that originates from individual steps in the SDH construction (gauged) and in the regionalization approach (ungauged). Finally, we propagate the uncertainty of the constructed SDHs in gauged catchments through SDH regionalization to ungauged catchments. This enables the quantification of the coupled uncertainty of SDHs that is composed of uncertainty from both the SDH construction and regionalization. Such SDHs with corresponding uncertainty bands should be provided to engineers and practitioners as reliable flood estimates (*Chowdhury and Stedinger*, 1991).

There is a lack of uniform terminology and a general disagreement about appropriate methodologies for uncertainty quantification in hydrological applications (*Nearing et al.*, 2016). We will adopt an uncertainty definition often used in flood frequency analysis, where uncertainty is expressed as the variability of the design value under consideration. *Serinaldi* (2013) stated that complementing accurate point estimates with realistic confidence intervals (CIs), which clearly highlight the lack of information, is probably the most correct approach to communicate results of hydrological frequency analyses.

The choice of the uncertainty estimation method often depends on the sources of uncertainty considered. The influence of a model choice can be best quantified through comparing a number of models (*Kidson and Richards*, 2005). On the contrary, sampling uncertainty related to parameter estimation is usually either assessed via the distribution of maximum likelihood (ML) estimators or via resampling approaches (*Beven and Hall*, 2014). A resampling approach often used is bootstrapping which involves randomly selecting data points, with replacement, from the original sample and then estimating the extreme flow quantile from each of the resampled data sets (*Efron and Tibshirani*, 1993; *Burn*, 2003; *Hall et al.*, 2004; *Wasserman*, 2006; *Meylan et al.*, 2012).

The presence of several uncertainty sources requires their joint consideration. The errors of the various sources are usually not independent from each other and are therefore not necessarily additive (*Montanari and Koutsoyiannis*, 2012; *Sikorska and Renard*, 2017). Hence, the total uncertainty is usually not necessarily equal to the sum of its contributing sources. Moreover, most of the models or design procedures used in hydro-systems engineering and analysis are nonlinear and highly complex. This prohibits an analytical derivation of the probability distribution of the model outputs. Engineers therefore frequently resort to methods that yield approximations for the statistical properties of model outputs that are subject to uncertainty (*Chang et al.*, 1994; *Tung and Yen*, 2005). A method often used to propagate uncertainties through a model chain is the Monte Carlo (MC) approach (*Beven et al.*, 2010). This approach is often used to assess the total uncertainty of a hydrological model output that involves observational, model, and parameter uncertainty. *Montanari and Koutsoyiannis* (2012) proposed to estimate the distribution of the output of a process-based hydrological model via multiple simulation runs by perturbing input data, parameters, and model output. In this study, we adopt this idea to a flood frequency model. We conduct a bootstrap experiment to assess the distribution of synthetic design hydrographs for a specific catchment while considering different uncertainty sources.

# 2 Methods

## 2.1 Synthetic design hydrographs

Synthetic design hydrographs (SDHs) describe not only the peak discharge of a flood but also its hydrograph volume and shape (*Brunner et al.*, 2017a).

### 2.1.1 Construction of synthetic design hydrographs in gauged catchments

*Brunner et al.* (2017a) proposed a method for the construction of SDHs in gauged catchments based on runoff data only. The method models the entire shape of the hydrograph using a probability density function (PDF), and estimates the design variable quantiles peak discharge and hydrograph volume considering their dependence. It consists of eight steps:

1. Flood sampling using a peak-over-threshold (POT) approach;
2. Baseflow separation using the recursive digital filter proposed by *Eckhardt* (2005) whose two parameters need to be estimated for each catchment;
3. Identification of the median hydrograph and its normalization. The median hydrograph is defined using the $h$-mode depth for functional data (*Cuevas et al.*, 2007). In its normalized form, we refer to it as the representative normalized hydrograph (RNH);
4. Fitting of a lognormal probability density function (PDF) (*Yue et al.*, 2002) to the RNH. The parameters of the PDF are computed as a function of the time to peak, the peak discharge, and the time base of the RNH (*Nadarajah*, 2007; *Rai et al.*, 2009);
5. Determination of marginal distributions of peak discharges and hydrograph volumes. The Generalized Pareto distribution (GPD) is used to model the marginal distribution of peak discharges and the Generalized extreme value (GEV) distribution to model the marginal distribution of the hydrograph volumes.
6. Dependence modeling between peak discharges and hydrograph volumes using the Joe copula (*Genest and Favre*, 2007; *Joe*, 2015) independently of the choice of their marginal distributions;
7. Estimation of the design variable quantiles peak discharge ($Q_T$) and hydrograph volume ($V_T$) for a chosen return period. Computation of the duration of the design event ($D_T$);
8. Composition of the design hydrograph using the shape of the hydrograph given by the PDF ($f(t)$), the design variable quantiles ($V_T$ and $D_T = V_T/Q_T$), and the baseflow ($B$) as described by
$$Q_T(t) = f(t)V_T/D_T + B. \tag{1}$$

We refer to the procedure described above as the *standard configuration* for obtaining an SDH. For a further, detailed description of the methodology, the reader is referred to *Brunner et al.* (2017a). The proposed methodology is generally applicable to any dataset of interest, however, the model assumptions made in Step 4 (choice of PDF), Step 5 (choice of marginal distributions), and Step 6 (choice of copula family) might need to be refined for a different dataset than the one used in this study.

The design flood hydrographs obtained using this method are composed of ten parameters, which we herein refer to as *SDH parameters*. Specifically, two parameters are needed to model the shape of the hydrograph defined by the lognormal PDF with a location and a scale parameter. Three parameters each (location, scale, and shape) are needed to model the marginal distributions of the design variables peak discharge and hydrograph volume. One parameter defines the dependence between these two variables and one parameter characterizes the proportion of baseflow with respect to the direct hydrograph.

### 2.1.2 Regionalization of synthetic design hydrographs to ungauged catchments

The SDH parameters can be transferred from gauged to ungauged catchments using methods based on the relation between catchment characteristics and model parameters, approaches based on spatial proximity, or on homogeneous regions. *Brunner et al.* (2017b) tested regionalization methods belonging to five categories: 1) linear regression models establishing a linear relationship between SDH parameters and catchment characteristics, 2) nonlinear regression models additionally exploiting nonlinear relationships between SDH parameters and catchment characteristics, 3) spatial approaches, 4) regional mean models for fixed regions formed according to catchment size and elevation, and 5) methods forming homogeneous regions to transfer the whole parameter set from similar catchments to the ungauged catchment. The methods of the second category were found to be most suitable for the transfer of SDH parameters to ungauged catchments. Among these methods, the nonlinear regression method, boosted regression trees, performed best. A boosted regression tree model is a linear combination of many regression trees (CART models). It can be thought of as a regression model where each term is a tree (*Elith et al.*, 2008; *Hofner et al.*, 2009). It builds successive trees in a stagewise procedure where new trees depend on previous trees. Only a proportion of the observations is selected at each step to fit the tree model in order to prevent from overfitting (*Friedman*, 2001; *Hastie et al.*, 2008). For a detailed overview and description of the methods tested, the reader is referred to *Brunner et al.* (2017b). For another application of boosted regression trees in the hydrological context, see *Tisseuil et al.* (2010).

## 2.2 Uncertainty assessment framework

Our uncertainty analysis takes into account several uncertainty sources inherent in both the construction process and in the regionalization process of SDHs. Therefore, we distinguish between these two processes and three levels of complexity for analyzing uncertainty (for an illustration see Figure 1). On a first level (A), we assess the individual uncertainty sources present in the SDH construction and the SDH regionalization separately (Section 2.2.1). On the second level (B), we assess the total uncertainty of the SDH construction and the total uncertainty of the SDH regionalization resulting from different sources separately (Section 2.2.2). On the third level (C), we propagate the uncertainty of the SDH construction through the regionalization process to couple it with the total uncertainty present in regionalization (Section 2.2.3). All these levels of the uncertainty analysis were based on bootstrap simulations. The basic principle was to construct a set of SDHs using various model configurations and to compare the characteristics of this set to the characteristics of an SDH obtained as a *best estimate* under the standard configuration (i.e., when no uncertainty is considered). The number of simulation runs was determined by a convergence analysis (*Beven et al.*, 2010). We tried to minimize computation time by keeping the number of runs low while ensuring stable estimates. This resulted in different numbers of simulation runs for different levels of the analysis (generally 100 to 1000 runs were sufficient, 2000 were used for the coupled uncertainty analysis. The exact numbers are indicated in the respective paragraphs). We did not consider observational uncertainty of runoff time series because neither sufficient information on runoff measurement error nor measurement technique was available for all the catchments in the dataset. Furthermore, we did not consider the influence of the choice of the parameter estimation method since *Dung et al.* (2015) found that confidence regions for bivariate quantiles are hardly influenced by the choice of the parameter estimation method.

### 2.2.1 Uncertainty due to individual sources

An investigation of the impact of individual uncertainty sources on the overall output uncertainty provides important information regarding the relative contribution of uncertainty sources to the overall uncertainty of the model output (here SDH) (*Tung and Yen*, 2005; *Sikorska et al.*, 2012).

We distinguished between three categories of uncertainty: 1) uncertainty due to a limited record length, 2) model uncertainty resulting from the choice of one model over another feasible model, and 3) sampling uncertainty resulting from estimating the model parameters based on an available flood sample that only approximates the characteristics of the underlying population. The first and the third category are closely related. While the first category explicitly considers the effect of the sample size, the third category keeps the sample size constant but considers that a slightly different sample could have been observed leading to different estimates. The steps in the SDH construction and regionalization procedure concerned with model and sampling uncertainty are listed in Table 1. It

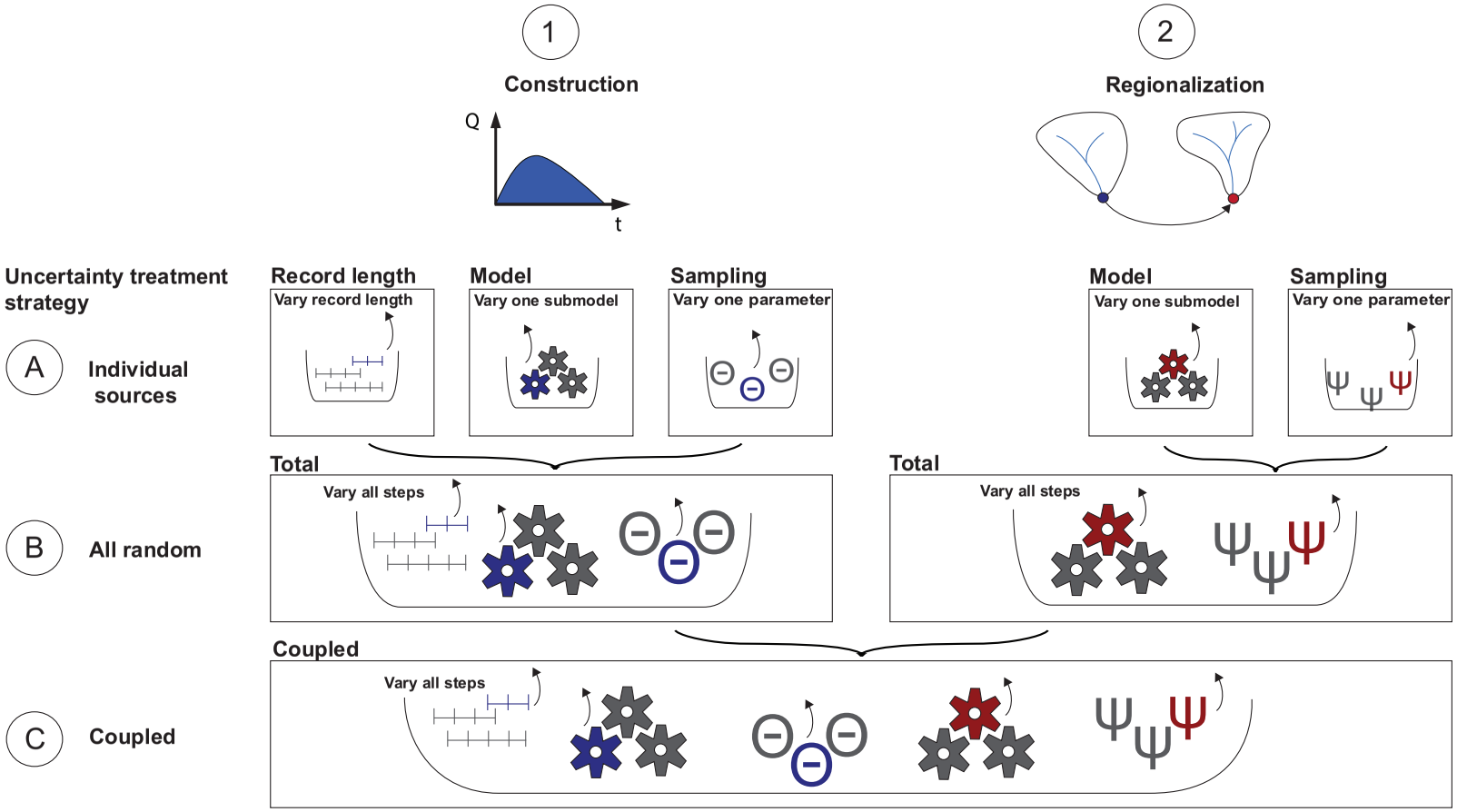



Figure 1: Illustration of the uncertainty framework proposed in this study. The uncertainty of constructed and regionalized SDHs was assessed on three levels of complexity: A) Uncertainty introduced by individual sources, specifically, record length, model choices, and parameter estimation; B) Total uncertainty of the constructed SDH and total uncertainty of the regionalized SDH resulting from the sources described in A; C) Coupled uncertainty of the SDH when construction uncertainty (steps 1A-1B) is propagated through regionalization (steps 2A-2B) onto the final regionalized SDH.

indicates which models were found to be most appropriate for each step (see Section 2.1.1) and which models would have also been feasible. Further, it highlights steps of parameter estimation which are subject to sampling uncertainty.

We focused on one uncertainty source at a time to assess the impact of the individual uncertainty sources on the estimated SDH. To do so, we constructed various SDHs using the standard model configuration while we either varied one model choice or considered the uncertainty of one parameter at a time. All the other model choices and parameters were fixed to the standard configuration (see Table 1: Model used). The set of SDHs obtained by these simulations gives an idea of the variability introduced by each source of uncertainty considered. The uncertainty assessment framework proposed is applicable to any return period sensible in the light of data availability. In our case study, we focused on a return period of 100 years because it is frequently used as protection goals for agricultural areas and settlements in Switzerland (*Camezind-Wildi*, 2005). A preliminary analysis has shown that the relative importance of different uncertainty sources is independent of the return period. In this work, we considered the following uncertainty sources (see also Table 1):

0. Record length: We computed SDHs to assess the effect of the sample size on the SDH estimates using subsamples of the original time series (*Botto et al.*, 2014; *Burn*, 2003). The samples were drawn without replacement from the original sample to exclude a potential effect of the chronological order of the events. We started with drawing a time series of 20 years and increased the length of the time series by 5 years at a time until the maximum available record length was reached. The minimum length was set to 20 years because a flood frequency analysis based on a shorter time series would provide unreliable flood estimates (*DVWK*, 1999). The stations with only 20 years of observations (6% of all stations) were excluded from this part of the analysis.

1. Flood sampling: We assessed model uncertainty due to flood sampling by constructing SDHs using flood events based on four flood sampling strategies: 1) peak-over-threshold approach (*Lang et al.*, 1999) sampling four events on average per year ($\mathrm{POT}_4$); 2) peak-over-threshold approach sampling two events on average per year ($\mathrm{POT}_2$); 3) annual peak maxima sampling ($\mathrm{AM}_Q$); 4) annual volume maxima sampling ($\mathrm{AM}_V$). To sample annual volume maxima, we computed the runoff volume over a running window of 72 hours, the maximum length of a frontal storm in Switzerland, and identified the window with the maximum volume per year.

2. Baseflow separation: Model uncertainty due to the baseflow separation method was not assessed since the two-parameter recursive digital filter proposed by *Eckhardt* (2005) was previously found to outperform alternative models for the application under consideration (*Brunner et al.*, 2017a). The recursive digital filter used requires the determination of two parameters: a recession coefficient $\alpha$ and a maximum baseflow coefficient $B_{\max}$. The uncertainty due to these two parameters was assessed via a parametric bootstrap by sampling the two parameters $N_{\mathrm{B}} = 100$ times from their distributions. $\alpha$ was sampled from a Weibull distribution fitted to the $\alpha$'s overall study catchments while $B_{\max}$ was sampled from a normal distribution fitted to the $B_{\max}$'s overall catchments. The Weibull and normal distributions were found to fit the data well based on the Kolmogorov–Smirnov goodness-of-fit test (level of significance of 0.05).

3. Normalization and identification of a representative normalized hydrograph: The identification of the representative normalized hydrograph (RNH) does, on the contrary to the normalization of the hydrograph, introduce uncertainty. The RNH was defined as the median hydrograph of the catchment under consideration. The $h$-mode depth was chosen to define the median hydrograph within a catchment among a set of four suitable definitions of depth functions for functional data. We considered the model uncertainty coming from the choice of one depth function over the others. For this, we constructed four SDHs using each of the suitable depth functions: $h$-mode depth, Fraiman–Muniz depth, random projection depth (*Cuevas et al.*, 2007), and band depth (*López-Pintado and Romo*, 2009). The definition of the RNH via the median does not involve parameter estimation, therefore, sampling uncertainty did not need to be considered.

4. Fitting of a probability density function to the RNH: The fitting of a PDF to the RNH to model the shape of the SDH introduces model uncertainty. One could fit another PDF to the RNH instead of the lognormal PDF, since the best PDF depends on the catchment. We chose the

Table 1: List of individual uncertainty sources inherent in the steps of the construction and regionalization of synthetic design hydrographs (Section 2.1.2). The methods used in each step are listed together with other feasible models and information on the parameters estimated.

| **Step** | **Model used (standard configuration)** | **Other feasible models** | **Estimated parameters** |
|---|---|---|---|
| 1 | Peak-over-threshold (POT) approach sampling four events per year on average | POT approach sampling two events per year on average<br>Annual maxima in terms of peak discharges<br>Annual maxima in terms of hydrograph volumes | - |
| 2 | Two parameter recursive digital filter (*Eckhardt*, 2005; *Collischonn and Fan*, 2013) | - | $\alpha$: recession coefficient estimated using linear regression; $B_{\max}$: maximum baseflow index which follows from $\alpha$ and discharge |
| 3 | Depth notion $h$-mode | Fraiman–Muniz depth<br>Band depth<br>random projection depth (*Cuevas et al.*, 2007) | - |
| 4 | lognormal PDF | normal PDF<br>Fréchet PDF<br>Weibull PDF<br>logistic PDF<br>gamma PDF<br>inverse gamma PDF<br>beta PDF | - |
| 5 | GPD for $Q_{\mathrm{p}}$ and GEV for $V$ | GEV for $Q_{\mathrm{p}}$<br>GPD for $V$ | Location, scale, and shape parameters of the GPD and the GEV estimated using the ML estimation |
| 6 | Use of Joe copula to model the dependence between $Q_{\mathrm{p}}$ and $V$ | Gumbel<br>Survival Clayton<br>Tawn | $\theta$: dependence parameter estimated based on pseudo-observations using maximum pseudo-likelihood estimation |
| 7 | Choice of the event with ML on the isoline in the bivariate space | Choice of other event on isoline | - |
| 8 | Mean event baseflow | - | Estimation of mean event baseflow |
| 9 | Nonlinear regression model boosting | Linear regression: lasso<br>Spatial: universal kriging<br>Regional mean: elevation zones<br>Formation of homogeneous regions: median parameter set from most similar catchment. | Parameters of the boosting model estimated by minimizing a loss function |

lognormal PDF because it showed the best average fit in terms of the Nash–Sutcliffe and Kling–Gupta efficiencies (*Gupta et al.*, 2009) when considering all catchments in the dataset. This model uncertainty was assessed by constructing SDHs using eight different PDFs: lognormal, normal, Fréchet, Weibull, logisitic, gamma, inverse gamma, and beta PDF. Sampling uncertainty due to the fitting of a PDF to the RNH was not considered since the estimation method chosen is based on an analytical expression (*Nadarajah*, 2007; *Rai et al.*, 2009). It expresses the parameters of the PDF in terms of the time to peak and the time base of the RNH instead of inferring the parameters via classical estimation approaches such as maximum likelihood or method of moments.

5. Determination of marginal distributions of peak discharges and hydrograph volumes: The determination of the marginal distributions for the two design variables peak discharge ($Q_\mathrm{p}$) and hydrograph volume ($V$) introduced both model and sampling uncertainty. Goodness-of-fit tests (Kolmogorov–Smirnov, Anderson–Darling, and upper tail Anderson–Darling test (*Chernobai et al.*, 2015)) of different model alternatives (Generalized Pareto (GPD), Generalized Extreme Value (GEV), Gumbel, generalized logistic, Pearson type III, normal, lognormal, exponential, and Weibull) showed that only the hypothesis of GPD and GEV distributions were not rejected in most of the catchments at $\alpha = 0.05$. The results of the goodness-of-fit tests suggested that the GPD and GEV models could both be used to model $Q_\mathrm{p}$ and $V$ and it was difficult to decide which distribution was a better predictor over another (*Beven and Hall*, 2014). We therefore assessed model uncertainty for the marginal distributions of peak discharges and hydrograph volumes by constructing two SDHs based on each of the two admissible marginal distributions. The parameters of the marginal distributions were estimated using ML estimation (*Held and Sabanés Bové*, 2014), which introduced sampling uncertainty. The ML approach assumes that the parameter estimates follow a parametric distribution. This distribution can be used to assess sampling uncertainty. We constructed $N_\mathrm{B} = 500$ SDHs by sampling parameters from a multivariate Normal distribution defined by the means and covariance matrices of the location, scale, and shape parameters of the marginal distributions. We only allowed positive location and scale parameters to be sampled because the location and scale parameters must be positive.

6. Dependence modeling between $Q_\mathrm{p}$ and $V$: The modeling of the dependence between $Q_\mathrm{p}$ and $V$ introduced both model and sampling uncertainty. Model uncertainty was introduced because several copula models were non-rejected in most catchments (*Brunner et al.*, 2017a). A statistical bootstrap procedure was applied to compute a $p$-value for the Cramér-von Mises statistic (*Genest et al.*, 2009) of twelve copulas (Independence, Gumbel, Clayton, Joe, Frank, AMH, normal, Student, Tawn, t-EV, Plackett, and Survival Clayton) (*Joe*, 2015). Only four copulas were not rejected at a level of significance of $\alpha = 0.05$ in most of the catchments: Joe, Gumbel, Survival Clayton, and Tawn. We used each of the non-rejected copulas to construct four SDHs. Sampling uncertainty was introduced because the copula parameter $\theta$ could only be estimated based on a sample as the underlying population was unknown. We assessed this uncertainty using a parametric bootstrap experiment in which we constructed $N_\mathrm{B} = 500$ SDHs sampling the $\theta$ from a lognormal distribution fitted to the $\theta$s of all the study catchments. The lognormal distribution was found to fit the $\theta$s well based on the Kolmogorov–Smirnov goodness-of-fit test. The $\theta$ of the Joe copula must be positive. Therefore, we only allowed the sampling of positive values.

7. Estimation of the design variable quantiles for peak discharge and hydrograph volume. The choice of the definition of the return period is defined by the problem at hand (*Serinaldi*, 2015; *Brunner et al.*, 2016) and the return period usually defined in national guidelines (*Requena et al.*, 2013). However, several design events have the "same" probability of occurrence in bivariate frequency analysis and thus lie on an isoline. Often, the most likely of these pairs of design variables is chosen for practical application (*Brunner et al.*, 2016). This choice, however, introduces uncertainty because one could choose another pair of design variables on the isoline instead of the one with the highest likelihood. This source of uncertainty was assessed by a bootstrap experiment with $N_\mathrm{B} = 1000$ SDHs sampling one point on the isoline according to the probability density function of the points on the isoline. Sampling according to the probability density function ensures that more points close to the most likely point and very few extreme points are sampled.

8. Composition of the design hydrograph: The SDH can be composed using Equation (1) based on the shape of the hydrograph given by the PDF and the magnitude of the event given by the design variable quantiles and the baseflow component. The uncertainty of the shape and of the magnitude has been considered above. Additional uncertainty may come from adding the baseflow component. We assessed these sources of uncertainty using a bootstrap experiment with $N_B = 100$ SDHs sampling a mean event baseflow index from a normal distribution defined by a mean and a standard deviation derived from the mean baseflow indices over all catchments. The normal distribution was used because it was found to fit the baseflow indices over all catchments well based on the Kolmogorov–Smirnov goodness-of-fit test.

9. Regionalization: The regionalization of the SDHs from gauged to ungauged catchments introduced both model and sampling uncertainty. *Brunner et al.* (2017b) found that the non-linear regresssion model boosting was most suitable to regionalize the magnitude of the SDH. However, they tested models of different categories among which several could be found that performed as well as the boosting model. Thus, the choice of one regionalization model may have introduced model uncertainty. This source of uncertainty was assessed by regionalizing the SDH parameters (best estimates obtained using the standard configuration) using five regionalization methods. The five models considered were found to have an acceptable performance in the regionalization of SDHs in the previous study (*Brunner et al.*, 2017b). These were:

    - Linear regression: lasso
    - Non-linear regression: boosted regression trees (boosting)
    - Kriging: universal kriging with catchment area as explanatory variable
    - Regional mean: elevation zones
    - Formation of homogeneous regions: median parameter set from the five most similar catchments determined using hydrological reasoning.

    The sampling uncertainty was assessed using a bootstrap experiment which focused on the most suitable regionalization method boosting (*Brunner et al.*, 2017b). We sampled with replacement $n = 163$ (original sample size) catchments from the catchment set. The boosting model was fitted for each SDH parameter using the data from the resampled catchment set. The ten SDH parameters were predicted for the original catchment set using the respective regionalization model. This procedure was repeated $N_B = 500$ times resulting in a separate distribution for each SDH parameter in each catchment. These catchment specific distributions were fitted by a normal distribution since parametric bootstrap is to be preferred over non-parametric bootstrap (*Kysely*, 2008). However, the normal distribution did not provide a good fit for the location and scale parameters of the marginal distributions of peak discharges and hydrograph volumes. These distributions were found to be bi-modal or skewed in most catchments (roughly 90 %) and were therefore difficult to represent by a theoretical distribution. For consistency reasons, we still used the normal distribution to represent the distribution of all SDH parameters in all catchments. Using a theoretical distribution allowed the dependency between the location and scale parameter of the marginal distributions for $Q_p$ and $V$ to be considered by sampling from a multivariate normal distribution. This would not have been possible when using the empirical distributions. The use of a normal distribution instead of the empirical distribution, however, implied that central values were overrepresented in samples supposed to show skewed or bi-modal rather than centered distributions. We constructed $N_B = 500$ SDHs for each catchment randomly sampling the ten SDH parameters from the catchment specific parameter distributions.

### 2.2.2 Total uncertainty

On the previous level of complexity of the uncertainty analysis, we considered the effect of each uncertainty source independently of the other uncertainty sources. Here, we combine all these sources and assess the resulting total uncertainty of the 1) SDH construction and of the 2) SDH regionalization. However, we did not yet consider here that the constructed SDHs used to fit a regionalization model are uncertain themselves. That means we considered the regionalization uncertainty independently of the construction uncertainty at this level of the analysis.

1. The total uncertainty coming from the SDH construction process was assessed for a return period of $T = 100$ years by an "all random" strategy. We use the term "all random" strategy for a bootstrap simulation in which model choices were not fixed (instead models were also sampled) and where sampling uncertainty was taken into account. At each step of the construction procedure, we randomly sampled one option from the model and/or parameter space (see Table 1 for a list of admissible models). The uncertainty sources considered in the "all random" strategy were all the same as already considered in the assessment of the uncertainty coming from the individual uncertainty sources (see Section 2.2.1). The procedures applied were also the same except that the sampling uncertainty due to the dependence modeling (Step 7) was extended from the Joe copula to the other three suitable copulas. The $\theta$s of the Gumbel and Survival Clayton copulas, as those of the Joe copula, could be described by a lognormal distribution. However, no theoretical distribution could be found for the $\theta$s of the Tawn copula. Therefore, we used the empirical distribution of the Tawn $\theta$s over all catchments. This strategy was repeated $N_{\mathrm{B}} = 1000$ times to construct 1000 SDHs.

2. The total uncertainty coming from the SDH regionalization was also assessed by a bootstrap simulation. In each iteration, a regionalization model was first sampled from the model space. Then, the model was fitted for each SDH parameter using the data from a resampled catchment dataset (with replacement) to also consider sampling uncertainty. The sample size was kept the same as the original sample size (i.e., number of catchments $n = 163$) to not confuse uncertainty coming from a limited sample size with sampling uncertainty due to not knowing the true population. However, the resampling with replacement introduced numerical problems (non-invertible covariance matrix) when using the regionalization model universal kriging. To solve this problem, we removed redundant stations (approximately one third of the samples) from the resampled catchment set when the regionalization model universal kriging was sampled in the first step. The fitted models were then used to make predictions for the SDH parameters of the catchments in the original catchment set. The bootstrap experiment was repeated $N_{\mathrm{B}} = 2000$ times (determined by a convergence analysis) to construct 2000 regionalized SDHs.

#### 2.2.3 Coupled uncertainty

In the previous step, we assumed that the regionalization uncertainty was independent from the construction uncertainty. On the third level of complexity, we combined both uncertainty sources ,i.e., the uncertainty in the construction of SDHs with the uncertainty in the regionalization of SDHs. To this end, the uncertainty of the constructed SDHs of a return period of $T = 100$ years was propagated through the regionalization process. We refer to this uncertainty as the coupled uncertainty of SDHs. To propagate the construction uncertainty, catchment specific distributions of the ten SDH parameters had to be defined. This was done by running the "all random" strategy already used for the assessment for the total construction uncertainty $N_{\mathrm{B}} = 1000$ times for each catchment. This provided 1000 values for each of the ten SDH parameters which were assumed as the empirical distribution of these parameters within a catchment. For the uncertainty propagation, we then sampled one value from each of the ten empirical parameter distributions for each catchment and fitted a randomly sampled regionalization model to the data of a resampled (with replacement) set of catchments. We then used the fitted model to predict the SDH parameters for the original catchment set. This procedure was repeated $N_{\mathrm{B}} = 2000$ times. This corresponds to the procedure already applied when assessing the total regionalization uncertainty. The only difference between the assessment of the total regionalization uncertainty and the coupled uncertainty assessment was that in the former the ten parameters in each catchment were fixed whereas in the latter the ten SDH parameters for each catchment were sampled from catchment specific empirical distributions.

## 3 Case study

This uncertainty assessment study was performed using 163 Swiss catchments (for a map and a complete list of the catchments and their catchment characteristics see *Brunner et al.* (2017b)) with a wide range of catchment characteristics and flood behaviors. The catchments selected have hourly flow series of at least 20 years duration ranging up to 53 years. The catchments' runoff is neither altered by regulated lakes upstream or inland canals nor by highly urbanized areas. The catchments

are small to medium-sized (6 to 1800 km$^2$), situated between 300 and 2800 m.a.s.l. with respect to mean elevation, and have no or only a small percentage of areas with glaciers.

# 4 Results

## 4.1 Quantification of uncertainty via median absolute relative error

Each step of the uncertainty analysis provided a set of SDHs that were obtained by varying one or several model choices and/or considering sampling uncertainty. These sets were compared to the *best estimate SDH* obtained when using the standard configuration (i.e., without considering any uncertainty). To this end, we computed the absolute relative error of each SDH in the set when compared to the best estimate SDH. We summarized this information for each catchment by taking the median of the absolute relative errors of all SDHs in the set considered. We refer to this as the median absolute relative error ($E_{\mathrm{MAR}}$). The median was taken instead of the mean because it puts relatively little weight on very extreme errors compared to the mean and is therefore more robust. The $E_{\mathrm{MAR}}$s of all catchments together provided information on the variability of the uncertainty introduced by a certain uncertainty source across catchments. The $E_{\mathrm{MAR}}$ was used on all three levels of complexity of the uncertainty assessment framework. It was computed for four important hydrograph characteristics: the peak discharge ($Q_{\mathrm{p}}$), the hydrograph volume ($V$), the time to peak ($t_{\mathrm{p}}$), and the half-recession time ($t_{\mathrm{p05}}$). The half-recession time was defined as the time from the peak to where the hydrograph falls back to 0.5 times the peak discharge. The characteristics $Q_{\mathrm{p}}$ and $V$ describe the magnitude of the event while the characteristics $t_{\mathrm{p}}$ and $t_{\mathrm{p05}}$ describe the shape of the hydrograph.

The results from the three levels of complexity are summarized in Table 2 and presented in the next few sections.

## 4.2 Uncertainty due to individual sources

The $E_{\mathrm{MAR}}$ due to the individual uncertainty sources belonging to the categories record length, model, and sampling uncertainty are displayed in Figure 2 for a return period of $T = 100$ years. The median uncertainty of the individual uncertainty sources over all catchments ranged between 0% and roughly 30%. We look at each hydrograph characteristic in turn.

The $E_{\mathrm{MAR}}$ of $Q_{\mathrm{p}}$ was influenced by most sources of uncertainty considered. Among them, the choice of the marginal distribution of $Q_{\mathrm{p}}$ (median across sites $E_{\mathrm{MAR}} = 20\%$), the sampling uncertainty of this distribution (10%), the choice of the sampling strategy (8%), baseflow addition (8%), and the record length (3%) had the most prominent influence. Other sources of uncertainty, in particular, the ones related to the shape of the hydrograph, only slightly affected the $E_{\mathrm{MAR}}$ of $Q_{\mathrm{p}}$. The $E_{\mathrm{MAR}}$ of $V$ was also affected by most sources of uncertainty. The largest $E_{\mathrm{MAR}}$ came from the sampling uncertainty of the marginal distribution of $V$ (across sites $E_{\mathrm{MAR}} = 12\%$), the choice of the sampling strategy (12%), the choice of the marginal distribution of $V$ (12%), the baseflow separation (11%), and record length (median 9%). The $E_{\mathrm{MAR}}$ of $t_{\mathrm{p}}$ was, as the other characteristics, also influenced by most uncertainty sources. Among them, the choice of the PDF (median across sites $E_{\mathrm{MAR}} = 27\%$), the choice of the sampling strategy (20%), the record length (19%), the RNH definition (16%) had the largest influence. Similarly, the $E_{\mathrm{MAR}}$ of $t_{\mathrm{p05}}$ was most strongly influenced by the record length (median across sites $E_{\mathrm{MAR}} = 17\%$) and the choice of the sampling strategy (17%).

**Influence of sampling strategy** As shown above, the sampling strategy was one of the uncertainty sources most strongly affecting the $E_{\mathrm{MAR}}$ of all four hydrograph characteristics. We therefore have a closer look at this source of uncertainty. Each of the four sampling strategies defined its own flood sample, which formed the basis of the subsequent analysis. The four flood samples were quite different. On the one hand, they had different sample sizes ($\mathrm{POT}_4 > \mathrm{POT}_2 > \mathrm{AM}_Q = \mathrm{AM}_V$), on the other hand, they differed in terms of their dependence structure. The sample of the $\mathrm{POT}_4$ strategy was largest and did not reject only four copulas: Joe, Gumbel, Survival Clayton, and Tawn. All these models were able to model tail dependence which was found to be present in the data according to the upper tail dependence coefficient (*Poulin et al.*, 2007). The sample of the $\mathrm{POT}_2$ strategy was half as large as the one obtained by the $\mathrm{POT}_4$ strategy. The non-rejected copulas for this strategy were, in most of the catchments: Survival Clayton, Placket, Student, Independence, Joe, Gumbel, and Tawn.

Table 2: Summary of uncertainties across the three levels of complexity for the four hydrograph characteristics $Q_{\mathrm{p}}$, $V$, $t_{\mathrm{p}}$, and $Q_{\mathrm{p05}}$. The uncertainties are provided in the form of the 1st, 2nd, and 3rd quartile of the $E_{\mathrm{MAR}}$s over all catchments. The numbers were rounded to two decimal places.

| | $Q_{\mathrm{p}}$ | | | $V$ | | | $t_{\mathrm{p}}$ | | | $t_{\mathrm{p05}}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Uncertainty sources** | **1st** | **2nd** | **3rd** | **1st** | **2nd** | **3rd** | **1st** | **2nd** | **3rd** | **1st** | **2nd** | **3rd** |
| Sample size | 0 | 0.03 | 0.06 | 0.02 | 0.09 | 0.19 | 0 | 0.19 | 0.5 | 0 | 0.17 | 0.45 |
| Sampling strategy | 0.05 | 0.08 | 0.17 | 0.08 | 0.12 | 0.2 | 0.11 | 0.2 | 0.27 | 0.11 | 0.17 | 0.24 |
| RNH definition | 0 | 0 | 0 | 0.01 | 0.03 | 0.04 | 0.07 | 0.16 | 0.23 | 0.01 | 0.02 | 0.03 |
| PDF choice | 0 | 0 | 0 | 0.03 | 0.04 | 0.06 | 0.24 | 0.27 | 0.32 | 0.04 | 0.06 | 0.09 |
| Copula choice | 0 | 0.01 | 0.01 | 0.01 | 0.02 | 0.03 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.02 |
| Choice on isoline | 0.02 | 0.03 | 0.04 | 0.04 | 0.05 | 0.07 | 0.07 | 0.09 | 0.11 | 0.07 | 0.09 | 0.11 |
| Margin $Q_{\mathrm{p}}$ | 0.12 | 0.2 | 0.34 | 0 | 0 | 0 | 0.09 | 0.14 | 0.2 | 0.09 | 0.14 | 0.2 |
| Margin $V$ | 0 | 0 | 0 | 0.08 | 0.12 | 0.16 | 0.08 | 0.12 | 0.16 | 0.08 | 0.12 | 0.16 |
| Baseflow separation | 0.03 | 0.04 | 0.06 | 0.07 | 0.11 | 0.16 | 0.09 | 0.13 | 0.21 | 0.1 | 0.15 | 0.2 |
| Baseflow addition | 0.06 | 0.08 | 0.1 | 0.06 | 0.08 | 0.1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Copula parameter | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Margin $Q_{\mathrm{p}}$ parameter | 0.08 | 0.1 | 0.14 | 0 | 0 | 0 | 0.07 | 0.1 | 0.14 | 0.07 | 0.1 | 0.14 |
| Margin $V$ parameter | 0 | 0 | 0 | 0.13 | 0.15 | 0.19 | 0.13 | 0.15 | 0.19 | 0.13 | 0.15 | 0.19 |
| **Total construction** | 0.21 | 0.26 | 0.35 | 0.28 | 0.34 | 0.4 | 0.41 | 0.45 | 0.52 | 0.37 | 0.44 | 0.5 |
| Regionalization model | 0.23 | 0.36 | 0.63 | 0.21 | 0.34 | 0.76 | 0.12 | 0.21 | 0.34 | 0.19 | 0.28 | 0.41 |
| Regionalization sampling | 0.11 | 0.15 | 0.2 | 0.12 | 0.16 | 0.22 | 0.18 | 0.22 | 0.3 | 0.18 | 0.22 | 0.3 |
| **Total regionalization** | 0.26 | 0.4 | 0.67 | 0.29 | 0.43 | 0.68 | 0.14 | 0.25 | 0.41 | 0.2 | 0.3 | 0.44 |
| **Coupled** | 0.47 | 0.59 | 0.94 | 0.45 | 0.55 | 0.66 | 0.43 | 0.48 | 0.54 | 0.44 | 0.48 | 0.53 |

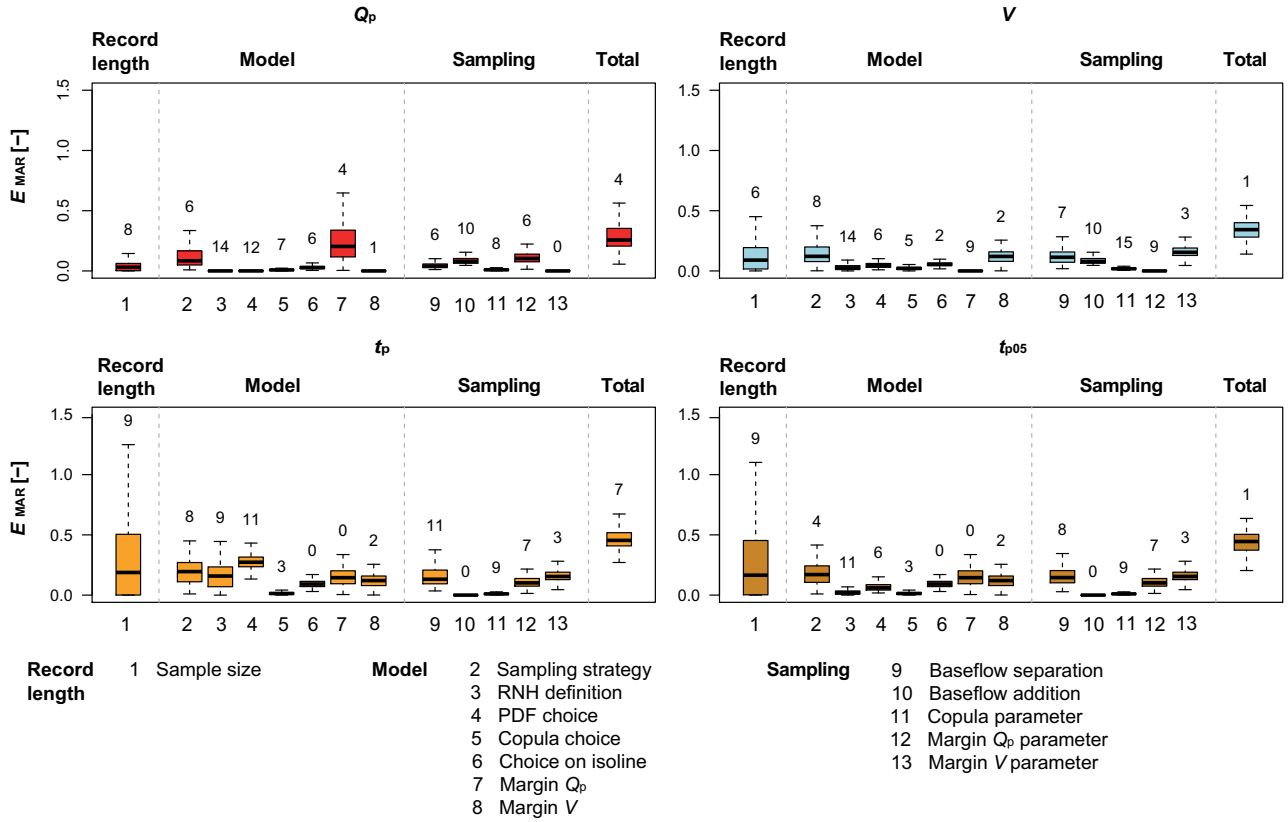



Figure 2: Median relative errors within the catchments of four hydrograph characteristics for an SDH of $T = 100$ years due to different uncertainty sources. The hydrograph characteristics assessed were peak discharge ($Q_\mathrm{p}$), hydrograph volume ($V$), time to peak ($t_\mathrm{p}$), and half-recession time ($t_\mathrm{p05}$). The whiskers extend to 1.5 times the interquartile range. Outliers are not displayed, however, their number among the 163 study catchments is given above the upper whisker of the boxplot.

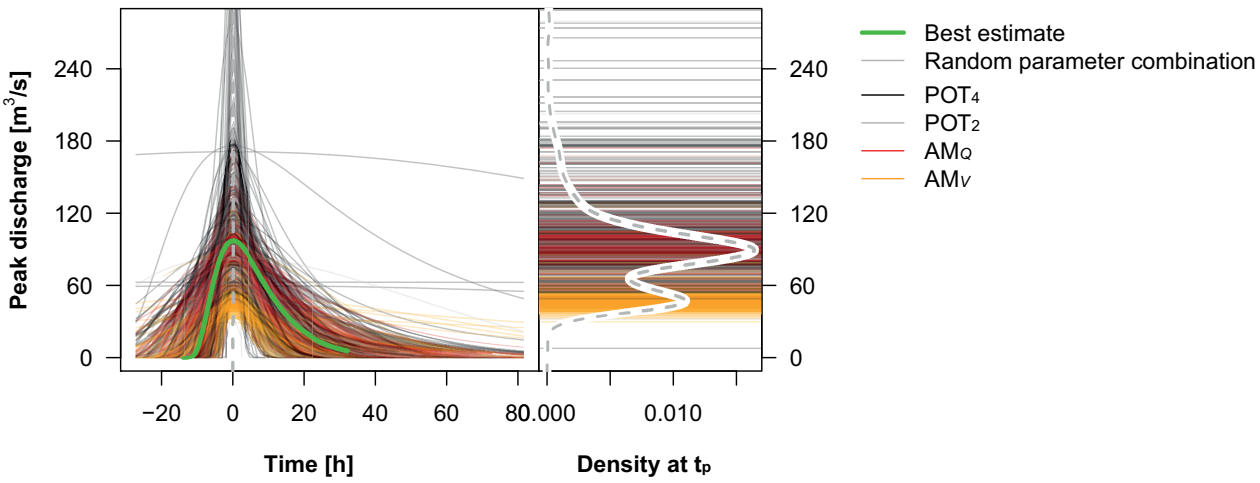



Figure 3: $N_{\mathrm{B}} = 1000$ SDHs obtained using the all random strategy for different sampling strategies for the Birse-Moutier catchment (left panel). SDHs obtained using a peak-over-threshold approach sampling four events per year on average ($\mathrm{POT}_4$) are colored in black, and two events per year on average ($\mathrm{POT}_2$) are colored in grey, SDHs obtained by annual maxima peak sampling ($\mathrm{AM}_Q$) are depicted in red and those obtained by annual maxima volume sampling ($\mathrm{AM}_V$) in orange. The best estimate SDH constructed using observed runoff data is given in green. The density of peak discharges is given in the right panel. Peaks are colored as in the left panel.

The sampling strategy $\mathrm{AM}_Q$ led to quite a small sample size and almost all of the copulas tested, including the independence copula, were non-rejected. The sampling strategy $\mathrm{AM}_V$ also led to a small sample and non-rejected the Survival Clayton, t-EV, normal, Frank, and Gumbel copulas in most of the catchments. The dependence between $Q_{\mathrm{p}}$ and $V$ assessed via the three correlation coefficients Pearson, Spearman, and Kendall was highest for the strategy $\mathrm{AM}_V$ and was comparable for the other three sampling strategies. The characteristics of the flood sample influenced the estimated SDHs, which is shown in Figure 3 for an example catchment. It shows that design hydrograph estimates derived based on the $\mathrm{AM}_V$ sample had generally a smaller magnitude than the estimates derived based on the flood samples obtained by the other strategies. Design flood estimates with very high peaks are mainly related to the sampling strategies $\mathrm{POT}_2$ and $\mathrm{POT}_4$.

**Regionalization uncertainty** Model and parameter choices both introduced uncertainty to the regionalized SDHs (see Figure 5). All four hydrograph characteristics were more affected by model uncertainty than by sampling uncertainty. The median $E_{\mathrm{MAR}}$ across all catchments related to model uncertainty lay around 30% for all hydrograph characteristics, while the median $E_{\mathrm{MAR}}$ related to sampling uncertainty was roughly 15% for peak discharges and hydrograph volumes and 20% for the characteristics related to the shape of the hydrograph. The characteristics describing the hydrograph shape ($t_{\mathrm{p}}$ and $t_{\mathrm{p05}}$) were more affected by sampling uncertainty than the characteristics describing the hydrograph magnitude ($Q_{\mathrm{p}}$ and $V$).

## 4.3 Total uncertainty

**Construction uncertainty** The $E_{\mathrm{MAR}}$ of the total construction uncertainty for the four hydrograph characteristics (for $T = 100$ years) was larger than each of the individual uncertainty sources (Figure 2), as we had expected. The largest median $E_{\mathrm{MAR}}$ (45%) over all catchments was found for the hydrograph characteristics related to the shape of the hydrograph ($t_{\mathrm{p}}$ and $t_{\mathrm{p05}}$) with an inter-quantile range of 40% to 50%. The smallest median $E_{\mathrm{MAR}}$ was found for $Q_{\mathrm{p}}$ (25%) with an interquartile

range spanning from 20% to 35%. The median $E_{\mathrm{MAR}}$ of $V$ over all catchments was 35% and the interquartile range spanned from 30% to 40%.

The SDHs obtained by the "all random" strategy are displayed in Figure 4 for three example catchments of different size: Langete-Huttwil (60 km$^2$), Mentue-Yvonand (105 km$^2$), and Birs-Münchenstein (911 km$^2$). The 90% confidence interval includes the best estimate SDH.

**Regionalization uncertainty** The $E_{\mathrm{MAR}}$ of the total regionalization uncertainty for the four hydrograph characteristics assessed for $T = 100$ years was as expected larger than the separate sampling and model uncertainties (Figure 5) even though the interquartile ranges are quite similar as those of the model uncertainties. The largest median $E_{\mathrm{MAR}}$ (40%) over all catchments was found for the hydrograph characteristics related to the magnitude of the hydrograph ($Q_{\mathrm{p}}$ and $V$) with an interquantile range of 25% to 65% and 30% to 70% respectively. The median $E_{\mathrm{MAR}}$ was lowest for $t_{\mathrm{p}}$ (25%) with an interquartile range spanning from 15% to 40%. The median $E_{\mathrm{MAR}}$ of $t_{\mathrm{p05}}$ was 30% and the interquartile range spanned from 20% to 45%.

## 4.4 Coupled uncertainty

The median $E_{\mathrm{MAR}}$ of the four catchment characteristics $Q_{\mathrm{p}}$, $V$, $t_{\mathrm{p}}$, and $t_{\mathrm{p05}}$ assessed via the coupled uncertainty strategy lay around 50% when looking at the median over all catchments (Figure 6). It was slightly lower for the characteristics related to the shape of the hydrograph than for those related to the magnitude. The $E_{\mathrm{MAR}}$ of the characteristics related to the magnitude showed a higher variability across the catchments. This variability was most pronounced for $Q_{\mathrm{p}}$.

# 5 Discussion

## 5.1 Uncertainty due to individual uncertainty sources

### 5.1.1 Record length

We showed that the record length is a major source of uncertainty in design hydrograph construction for all four hydrograph characteristics considered. *Serinaldi* (2013) stated that many joint distributions and copulas can be fitted to small samples because goodness-of-fit tests cannot discriminate between alternative models due to the lack of power. An increase in the length of relatively short samples could noticeably reduce the uncertainty of design variable quantiles (*Botto et al.*, 2016) and therefore, as shown in our study, the uncertainty of the whole design hydrograph. The problem of a limited sample size could for practical applications be partially overcome by temporal, spatial, or causal information expansion. Historical floods could be introduced to expand information in time, data from neighboring or similar stations could be used to expand data in space, and runoff processes could be considered for causal information expansion (*Blöschl et al.*, 2013).

### 5.1.2 Sampling strategy

Another major uncertainty source for all four studied hydrograph characteristics was the choice of the sampling strategy. It directly influences the sample size and the characteristics of the flood sample. In most catchments, the flood sample chosen by the strategy $\mathrm{AM}_V$ was characterized by lower peak discharges than the flood samples determined by the strategies $\mathrm{POT}_4$ and $\mathrm{AM}_Q$. This is the case because a high volume is not always related to a high peak discharge for all floods. *Brunner et al.* (2017a) have shown that the dependence between $Q_{\mathrm{p}}$ and $V$ is rather small for specific short duration flood-types such as flash floods and short-rain floods compared to long-rain floods (*Merz and Blöschl*, 2003; *Sikorska et al.*, 2015). From a safety point of view, the choice of the strategy $\mathrm{POT}_4$ seems to be judicious since the magnitude of the SDHs is not underestimated compared to the other strategies. The choice of the sampling strategy also influences the choice of a suitable copula. The sampling strategies leading to a small sample size ($\mathrm{AM}_Q$ and $\mathrm{AM}_V$) make the choice of a suitable copula inconclusive. Almost all copulas tested were not rejected because of the small power of the goodness-of-fit test in the case of small sample sizes (*Cullen and Frey*, 1999; *Genest et al.*, 2009). On the contrary, the choice of a sampling strategy with a larger sample size ($\mathrm{POT}_4$) is quite conclusive since only a few copulas are non-rejected. The choice of the sampling strategy $\mathrm{POT}_4$ therefore seems to be a suitable choice for the estimation of SDHs.

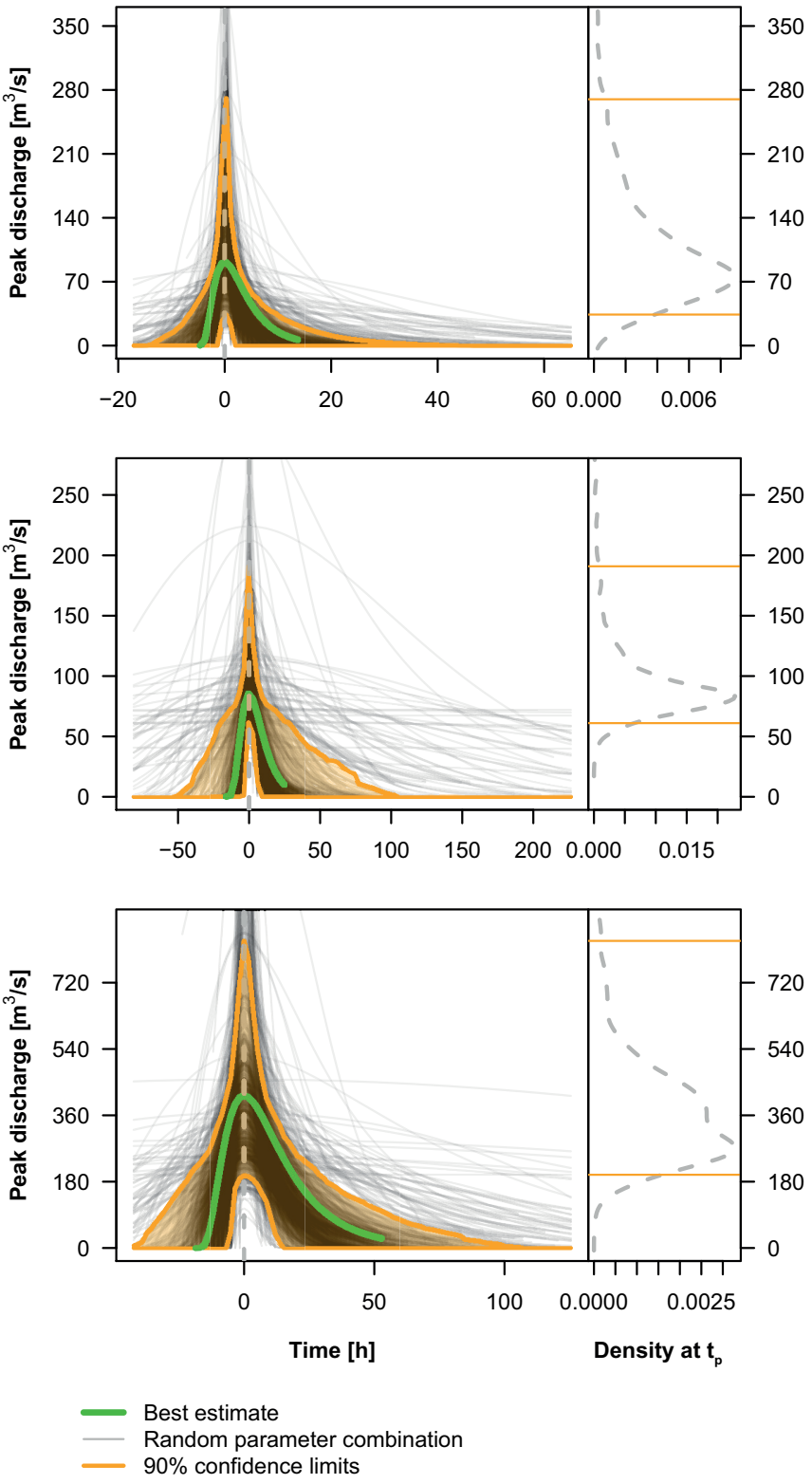



Figure 4: Total uncertainty of SDH construction assessed via the all random strategy for three catchments of different size: Langete-Huttwil (60km$^2$), Mentue-Yvonand (105km$^2$), and Birs-Münchenstein (911km$^2$) (from top to bottom). The best estimate obtained based on observed data (green) is shown together with the 90% confidence limits (orange) and the 1000 SDHs obtained by randomly picking model choices and parameters (grey). The orange shaded areas represent the 90% confidence regions. The right panel shows the density of peak discharges and the corresponding 90% confidence limits.

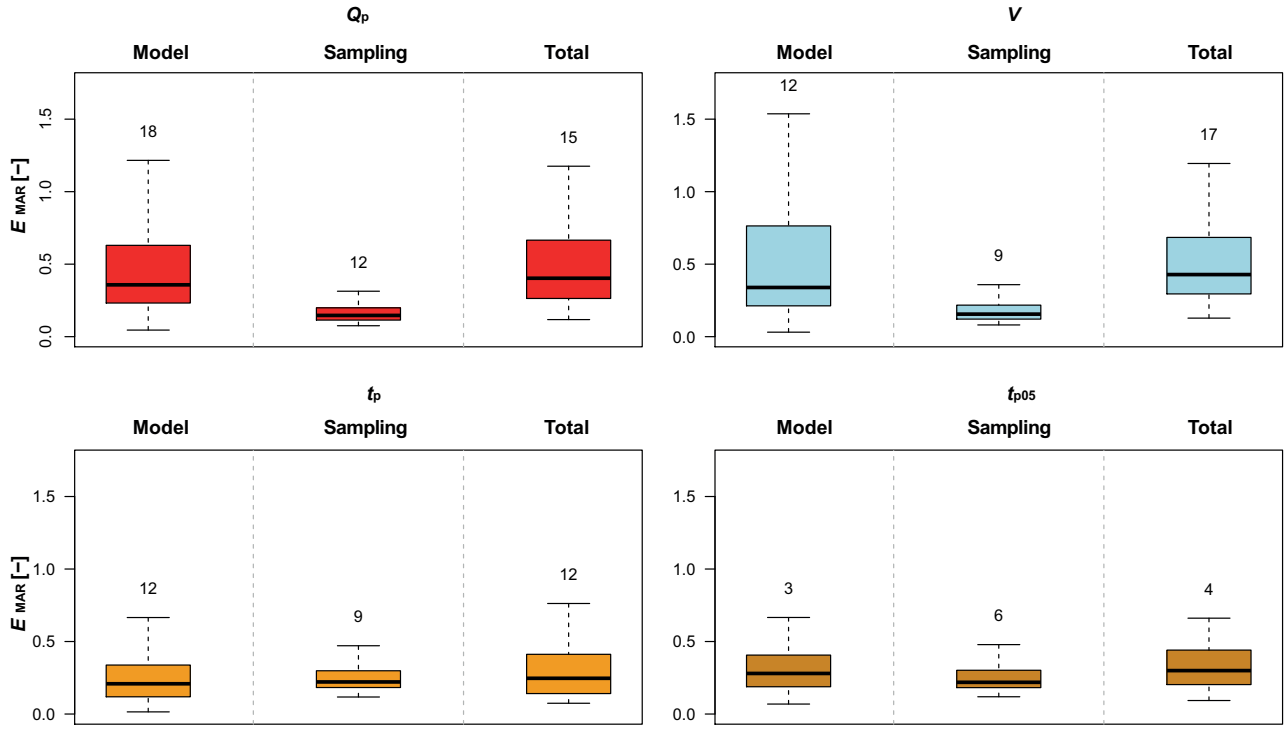



Figure 5: Regionalization uncertainty expressed as the median relative errors within the catchments of the four hydrograph characteristics: peak discharge ($Q_\mathrm{p}$), hydrograph volume ($V$), time to peak ($t_\mathrm{p}$), and half-recession time ($t_\mathrm{p05}$). Model and sampling uncertainty are given separately and combined (total). The whiskers extend to 1.5 times the interquartile range.
Outliers are not displayed, however, their number among the 163 study catchments is given above the upper whisker of the boxplot.

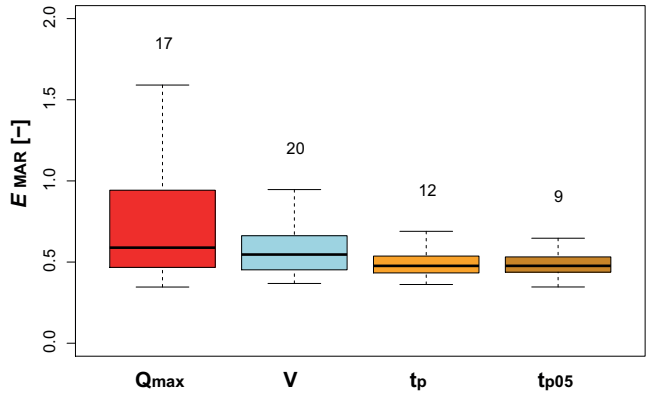



Figure 6: Uncertainty obtained by propagating construction uncertainty through regionalization. Uncertainty is expressed as the median relative errors within catchments of the four hydrograph characteristics: peak discharge ($Q_\mathrm{p}$), hydrograph volume ($V$), time to peak ($t_\mathrm{p}$), and half-recession time ($t_\mathrm{p05}$). The whiskers extend to 1.5 times the interquartile range. Outliers are not displayed, however, their number among the 163 study catchments is given above the upper whisker of the boxplot.

### 5.1.3 Model uncertainty

Besides record length and sampling strategy, various other uncertainty sources influenced the four analyzed characteristics of a design hydrograph. The magnitude of the event was largely influenced by the choice of the marginal distributions for peak discharges and hydrograph volumes and their corresponding sampling uncertainty. Still, modelers are faced with the problem of determining one model to be applied to a catchment for a particular modeling task (*Marshall et al.*, 2005; *Ajami et al.*, 2007; *Kuczera et al.*, 2010). *Beven and Hall* (2014) and *Kite* (1975) stated that the choice of one out of several models that are not rejected by statistical goodness-of-fit tests might not be straightforward but does not matter in the range of the observed data. However, the choice of one model over another might matter a lot when extrapolating beyond the range of the data (*Hosking and Wallis*, 1997). Sampling uncertainty related to the estimation of the parameters of the marginal distributions has been found to be important in univariate (*Lamb and Kay*, 2004; *Hailegeorgis and Alfredsen*, 2017) and in bivariate quantile estimation (*Serinaldi*, 2013; *Dung et al.*, 2015). Unlike *Serinaldi* (2009), we did not find that model uncertainty is smaller than sampling uncertainty. The relative contribution of error due to model choice and sampling error most likely depends again on the record length. For large samples, the standard deviation of the estimate becomes small in comparison to the bias caused by the wrong distribution choice (*Strupczewski et al.*, 2002).

The magnitude of the flood event is only slightly affected by the choice of one of the non-rejected copulas. This is in accordance with findings by *Xu et al.* (2010) who found that the uncertainty originating from method selection was largest and the uncertainty caused by ignoring the dependence was smallest. However, this might be different if one is interested in estimates for higher return periods, where the ability of a copula to correctly describe the tail dependence in the data might be crucial (*Poulin et al.*, 2007). Our results highlight the importance of a sufficiently large sample as obtained by $\mathrm{POT}_4$ compared to the annual maxima sampling strategies. The choice of a larger sample narrows down the number of admissible copulas. It prevents from having to choose a dependence structure based on a goodness-of-fit test with limited power when several probabilistic models cannot be rejected despite their poor fit to the data with the limited size of an annual maxima sample (*Cullen and Frey*, 1999). The uncertainty related to the choice of a design event on the isoline of equally likely events is also rather small compared to the other uncertainty sources considered. This is important since this uncertainty source depends on model selection and not on the sample size and is therefore irreducible.

Contrary to the magnitude, the characteristics related to the hydrograph shape were more influenced by the definition of the representative normalized hydrograph and the choice of a probability density function to model the shape of the hydrograph. The importance of the different uncertainty sources differed for the four hydrograph characteristics. Consequently, it is important to consider both uncertainty sources related to the magnitude of the event and sources related to the shape of the event when assessing the uncertainty of a design hydrograph.

### 5.1.4 Regionalization method

The relative importance of model and sampling uncertainty due to regionalization differed for the hydrograph characteristics related to the flood magnitude ($Q_\mathrm{p}$ and $V$) and those related to the hydrograph shape ($t_\mathrm{p}$ and $t_{\mathrm{p}05}$). Model uncertainty was higher than sampling uncertainty for most hydrograph characteristics. However, sampling uncertainty seemed to be quite important in the regionalization of the hydrograph shape.

## 5.2 Total uncertainty

The total uncertainty of constructed hydrographs was higher than each of the individual uncertainty sources considered. However, it was lower than the sum of all the individual uncertainty sources. In the case of regionalization, the spread of the $E_\mathrm{MAR}$s referring to model uncertainty were comparable to those referring to total uncertainty. This implies that model uncertainty alone introduces a lot of uncertainty to SDH regionalization and the effect of sampling uncertainty is negligible. The non-additive property of individual uncertainty sources highlights the need of considering all sources jointly when the total uncertainty is of interest. A simple adding of individual sources would result in highly overestimated and unrealistic uncertainties. The assessment of individual sources may be

thus only used to compare their relative contributions and to indicate which uncertainty part could be reduced with the highest benefit.

## 5.3 Coupled uncertainty

The coupled SDH uncertainty considering construction and regionalization uncertainty amounted to a bit more than 50% depending on the hydrograph characteristic considered. It was slightly higher for the hydrograph characteristics related to the magnitude of the event when looking at a return period of $T = 100$ years. For lower return periods, it would most likely be the other way around since it would be in the interpolation range of the bivariate extreme value distribution rather than the extrapolation range. The higher uncertainty for magnitude-related hydrograph characteristics compared well with the total uncertainties obtained for the regionalization. On the contrary, the total uncertainty of the constructed SDHs was higher for the hydrograph characteristics related to the hydrograph shape than for those related to the hydrograph magnitude. The coupled uncertainty was slightly higher than the total uncertainty in gauged catchments. This suggests that already SDH construction contributes a big part of the uncertainty and regionalization just adds a little uncertainty. Yet, these two uncertainty sources are not necessarily additive. The coupled uncertainty was lower than the sum of the total construction and the total regionalization uncertainty which implies that these have to be considered jointly, as it was done here, to get realistic uncertainty bands.

## 5.4 Value of the proposed uncertainty framework

The three level uncertainty framework proposed here allows the quantification of the uncertainty of synthetic design hydrographs as they would be provided to engineers or practitioners. It is therefore the first step towards communicating the uncertainty of design estimates to practitioners. SDHs complemented with uncertainty bands allow more informed decisions than using SDHs without uncertainty information (*Hall*, 2003). These simulation results obtained could be provided to the practitioner as a set of values comprising the best estimate, the first and third quantile of the simulation results and the maximum simulated SDH. Depending on the question of interest, the practitioner could choose the value to work with. If safety is crucial, a more conservative estimate could be used. If cost-efficiency is crucial, a smaller estimate could be chosen. Ideally, an ensemble-based approach could be adopted. An uncertainty assessment based on three levels of complexity further points out which individual uncertainty sources and which steps in the analysis offer the largest potential for improvement. This information can be used to decide which avenues to take for future research or in depth analysis. Our results suggest that uncertainty could be successively reduced if more data became available or by using alternative data.

We have also shown that the choice of a flood sampling strategy is quite an important source of uncertainty. We showed that using an annual maxima sampling strategy might neglect important information. *Singh et al.* (2005) also found that floods within a year may not be adequately represented by their maximum and suggested the use of two independent subpopulations: snowfall and rainfall generated floods. Such a conscious selection of more than one flood event per year might be an alternative to the selection of flood events using a peak-over-threshold approach without considering flood generating mechanisms. A subdivision into floods generated by different processes, as proposed by *Sikorska et al.* (2015) and *Merz and Blöschl* (2003) would further allow for a representation of different phenomena by different statistical distributions (*Shu and Ouarda*, 2008). Working with subpopulations would also allow the representation of different hydrograph shapes within a catchment. This would give a better impression of hydrograph shape variability within a catchment due to different processes and to different rainfall inputs than working with only one catchment specific hydrograph shape. The construction of flood type specific hydrographs as proposed by *Brunner et al.* (2017a) goes into this direction.

## 5.5 Limitations and Perspectives

The uncertainty assessment framework proposed here is generally applicable to any return period of interest. However, several choices need to be made dependent on the dataset: the choice of a set of distribution functions to fit the peak discharges and hydrograph volumes, the choice of a set of suitable copula models to model the dependence between these two variables, and the choice of

a set of PDFs to model the hydrograph shapes to be sampled from the simulations. Conducting a simulation study on several catchments is computationally quite expensive, however, such an analysis is usually only done once.

The uncertainty analysis presented here considered most uncertainty sources affecting synthetic design hydrographs. One possibly important uncertainty source was not considered here: observational uncertainty due to measurement errors and due to the conversion of water level to discharge via a rating curve approach. The findings on the importance of this source of uncertainty differ quite a bit between various authors. *Apel et al.* (2004) showed that the stage-discharge relationships are relatively less important in the uncertainty analysis while *Kundzewicz* (2002) claimed that uncertainty resulting from the need to extrapolate the rating curve to extreme values, where no direct runoff measurements exist, was considerable. We did not consider this uncertainty source because the uncertainty of the rating curve was not known for the whole set of catchments. Therefore, it could not be incorporated into the uncertainty assessment framework proposed here. Still, rating curve uncertainty might be substantial especially when looking at extreme events (*Di Baldassarre and Montanari*, 2009) as it will propagate through the whole SDH construction process. In addition, the uncertainty due to the rating curve is catchment dependent and information on this uncertainty is often limited and case-specific (*Di Baldassarre and Montanari*, 2009; *Sikorska et al.*, 2013), which makes the generalization of the rating curve uncertainty difficult. The effect of the rating curve uncertainty could therefore be potentially assessed only for a few catchments in our dataset where enough information on the rating curve and its uncertainty is available.

# 6 Conclusions

Synthetic design hydrographs are inherently uncertain. They are affected by various uncertainty sources comprising a limited record length, model, and sampling uncertainties. This uncertainty needs to be assessed and communicated to the practitioner to ensure reliable flood estimates. In this work, we proposed a framework for assessing the uncertainty of such synthetic design hydrographs which consists of three levels of complexity using an extensive set of 163 Swiss catchments. First, we assessed the influence of individual sources of uncertainty on the estimation of synthetic design hydrographs. Second, we quantified the total uncertainty of constructed and regionalized synthetic design hydrographs considered separately. Third, we quantified the coupled uncertainty of a regionalized hydrograph when considering that already constructed hydrographs are uncertain. We identified the record length and the choice of the sampling strategy as having the strongest effect on the uncertainty of a synthetic design hydrograph characterized by peak discharge, hydrograph volume, and hydrograph shape. The magnitude of the design hydrograph was further strongly affected by the choice of the marginal distributions for peak discharges and hydrograph volumes. The shape of the hydrograph, however, was more affected by the definition of a representative hydrograph shape, the choice of the probability density function used to model this shape, and baseflow separation. The total uncertainty of design hydrographs constructed based on observed runoff data differed between the catchments analyzed and had a median of roughly 25% for $Q_{\mathrm{p}}$, 35% for $V$, 45% for $t_{\mathrm{p}}$ and $t_{\mathrm{p05}}$. The median $E_{\mathrm{MAR}}$s for hydrographs obtained by regionalization assuming that the data used for regionalization (constructed hydrographs) is known lie around 40% for $Q_{\mathrm{p}}$ and $V$, around 25% for $t_{\mathrm{p}}$ and around 30% for $t_{\mathrm{p05}}$. The coupled uncertainty of synthetic design hydrographs for ungauged catchments lay around 50% but differed quite a bit between catchments especially in terms of peak discharges. We also demonstrated that the uncertainty framework provides insights into promising avenues for future research. The uncertainty of synthetic design hydrographs could be most effectively reduced by enlarging the sample size by successively collecting more data or by considering alternative information such as historical flood data or regional data. Alternatively, we could aim at more adequately describing the variability of hydrograph types. Flood estimates complemented with uncertainty bands computed via the uncertainty assessment framework proposed in this study should allow the engineer to make profound decisions based on a cost-benefit analysis.

# Acknowledgement

We thank the Federal Office for the Environment (FOEN) for funding the project (contract 13.0028.KP / M285-0623) and for providing runoff measurement data. We also thank MeteoSwiss for pro-

viding precipitation data. The data used in this study is available upon order from the FOEN and MeteoSwiss. For the hydrological data of the federal stations, the order form under `http://www.bafu.admin.ch/wasser/13462/13494/15076/index.html?lang=de` can be used. The hydrological data of the cantonal stations can be ordered from the respective cantons. The meteorological data can be ordered via `https://shop.meteoswiss.ch/index.html`. We thank the three anonymous reviewers for their detailed and constructive feedback which helped us to improve the previous version of the manuscript.

# References

Ajami, N. K., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resources Research*, *43*, 1–19, doi:10.1029/2005WR004745.

Apel, H., A. H. Thieken, B. Merz, and G. Blöschl (2004), Flood risk assessment and associated uncertainty, *Natural Hazards and Earth System Science*, *4*(2), 295–308, doi:10.5194/nhess-4-295-2004.

Beven, K., and J. Hall (2014), *Applied uncertainty analysis for flood risk management*, 672 pp., Imperial College Press, London.

Beven, K., D. Leedal, and R. Alcock (2010), Uncertainty and Good Practice in Hydrological Prediction, *VATTEN*, *66*(3-4), 159–163.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (2013), *Runoff prediction in ungauged basins*, 465 pp., Cambridge University Press, Cambridge.

Botto, A., D. Ganora, F. Laio, and P. Claps (2014), Uncertainty compliant design flood estimation, *Water Resources Research*, *50*, 4242–4253, doi:10.1002/2013WR014981.

Botto, A., D. Ganora, P. Claps, and F. Laio (2016), Technical Note: Design flood under hydrological uncertainty, *Hydrology and Earth System Sciences Discussions*, p. in review, doi:10.5194/hess-2016-637.

Brunner, M. I., J. Seibert, and A.-C. Favre (2016), Bivariate return periods and their importance for flood peak and volume estimation, *Wire's Water*, *3*, 819–833, doi:10.1002/wat2.1173.

Brunner, M. I., D. Viviroli, A. E. Sikorska, O. Vannier, A.-C. Favre, and J. Seibert (2017a), Flood type specific construction of synthetic design hydrographs, *Water Resources Research*, *53*, doi:10.1002/2016WR019535.

Brunner, M. I., R. Furrer, A. E. Sikorska, D. Viviroli, J. Seibert, and A.-C. Favre (2017b), Synthetic design hydrographs for ungauged catchments: A comparison of regionalization methods, *Stochastic Environmental Research and Risk Assessment*, p. under review.

Burn, D. H. (2003), The use of resampling for estimating confidence intervals for single site and pooled frequency analysis, *Hydrological Sciences Journal*, *48*(1), 25–38, doi:DOI10.1623/hysj.48.1.25.43485.

Camezind-Wildi, R. (2005), Empfehlung Raumplanung und Naturgefahren, *Tech. rep.*, Bundesamt für Raumentwicklung, Bundesamt für Wasser und Geologie, Bundesamt für Umwelt, Wald und Landschaft, Bern.

Chang, C. H., Y. K. Tung, and J. C. Yang (1994), Monte Carlo simulation for correlated variables with marginal distributions, *Journal of Hydraulic Engineering*, *120*(3), 313–331, doi:10.1061/(ASCE)0733-9429(1994)120:3(313).

Chernobai, A., S. T. Rachev, and F. J. Fabozzi (2015), Composite Goodness-of-Fit Tests for Left-Truncated Loss Samples, in *Handbook of financial econometrics and statistics*, edited by C.-F. Lee and J. Lee, chap. 20, pp. 575–596, Springer Science+Business Media, New York.

Chowdhury, J. U., and J. R. Stedinger (1991), Confidence interval for design floods with estimated skew coefficient, *Journal of Hydraulic Engineering*, *117*(7), 811–831, doi:10.1061/(ASCE)0733-9429(1991)117:7(811).

Collischonn, W., and F. M. Fan (2013), Defining parameters for Eckhardt's digital baseflow filter, *Hydrological Processes*, *27*, 2614–2622.

Cuevas, A., M. Febrero, and R. Fraiman (2007), Robust estimation and classification for functional data via projection-based depth notions, *Computational Statistics*, *22*(3), 481–496, doi:10.1007/s00180-007-0053-0.

Cullen, A. C., and H. C. Frey (1999), *Probabilistic techniques in exposure assessment*, 335 pp., Society for Risk Analysis, New York.

Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall (2012), Merkblatt DWA-M 552, *Tech. rep.*, DWA, Hennef, Germany.

Di Baldassarre, G., and A. Montanari (2009), Uncertainty in river discharge observations: a quantitative analysis, *Hydrol. Earth Syst. Sci*, *13*, 913–921, doi:10.5194/hessd-6-39-2009.

Dung, N. V., B. Merz, A. Bárdossy, and H. Apel (2015), Handling uncertainty in bivariate quantile estimation - An application to flood hazard analysis in the Mekong Delta, *Journal of Hydrology*, *527*, 704–717, doi:10.1016/j.jhydrol.2015.05.033.

DVWK (1999), Statistische Analyse von Hochwasserabflüssen, *Tech. rep.*, Deutscher Verband für Wasserwirtschaft, Bonn.

Eckhardt, K. (2005), How to construct recursive digital filters for baseflow separation, *Hydrological Processes*, *19*, 507–515, doi:10.1002/hyp.5675.

Efron, B., and R. Tibshirani (1993), *An introduction to the bootstrap*, 436 pp., Chapman & Hall/CRC, Dordrecht.

Elith, J., J. R. Leathwick, and T. Hastie (2008), A working guide to boosted regression trees, *Journal of Animal Ecology*, *77*, 802–813, doi:10.1111/j.1365-2656.2008.01390.x.

Friedman, J. H. (2001), Greedy function approximation: a gradient boosting machine, *The Annals of Statistics*, *29*(5), 1189–1232.

Genest, C., and A.-C. Favre (2007), Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, *12*(4), 347–367, doi:10.1061/(ASCE)1084-0699(2007)12:4(347).

Genest, C., B. Rémillard, and D. Beaudoin (2009), Goodness-of-fit tests for copulas: A review and a power study, *Insurance: Mathematics and Economics*, *44*, 199–213, doi:10.1016/j.insmatheco.2007.10.005.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, *377*, 80–91, doi:http://dx.doi.org/10.1016/j.jhydrol.2009.08.003.

Hailegeorgis, T. T., and K. Alfredsen (2017), Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway, *Journal of Hydrology: Regional Studies Journal of Hydrology*, *9*, 104–126, doi:10.1016/j.ejrh.2016.11.004.

Hall, J., and D. Solomatine (2008), A framework for uncertainty analysis in flood risk management decisions, *International Journal of River Basin Management*, *6*(2), 85–98, doi:10.1080/15715124.2008.9635339.

Hall, J. W. (2003), Handling uncertainty in the hydroinformatic process, *Journal of Hydroinformatics*, *5*(4), 215–232.

Hall, M. J., H. F. P. van den Boogaard, R. C. Fernando, and a. E. Mynett (2004), The construction of confidence intervals for frequency analysis using resampling techniques, *Hydrology and Earth System Sciences*, *8*(2), 235–246, doi:10.5194/hess-8-235-2004.

Hastie, T., R. Tibshirani, and J. Friedman (2008), *The elements of statistical learning*, Springer series in statistics, 745 pp., Springer, Stanford, California.

Held, L., and D. Sabanés Bové (2014), *Applied statistical inference. Likelihood and Bayes*, 376 pp., Springer, Berlin-Heidelberg.

Hofner, B., A. Mayr, N. Robinzonov, and M. Schmid (2009), Model-based Boosting in R. A Hands-on Tutorial Using the R Package mboost, *Tech. rep.*, Department of statistics. University of Munich, Munich.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional frequency analysis*, 238 pp., Cambridge University Press, Cambridge, doi:10.1017/CBO9780511529443.

Joe, H. (2015), *Dependence Modeling with Copulas*, 479 pp., CRC Press. Taylor & Francis Group, Boca Raton.

Juston, J. M., A. Kauffeldt, B. Q. Montano, J. Seibert, K. J. Beven, and I. K. Westerberg (2013), Smiling in the rain: Seven reasons to be positive about uncertainty in hydrological modelling, *Hydrological Processes*, *27*(7), 1117–1122, doi:10.1002/hyp.9625.

Kidson, R., and K. S. Richards (2005), Flood frequency analysis: assumptions and alternatives, *Progress in Physical Geography*, *29*(3), 392–410, doi:10.1191/0309133305pp454ra.

Kite, G. W. (1975), Confidence limits for design events, *Water Resources Research*, *11*(1), 48–53, doi:10.1029/WR011i001p00048.

Klein, B., M. Pahlow, Y. Hundecha, and A. Schumann (2010), Probability analysis of hydrological loads for the design of flood control systems using copulas, *Journal of Hydrologic Engineering*, *15*(5), doi:10.1061/(ASCE)HE.1943-5584.0000204.

Koutsoyiannis, D. (2014), Reconciling hydrology with engineering, *Hydrology Research*, *45*(September), 2–22, doi:10.2166/nh.2013.092.

Kuczera, G., B. Renard, M. Thyer, and D. Kavetski (2010), There are no hydrological monsters, just models and observations with large uncertainties!, *Hydrological Sciences Journal*, *55*(6), 980–991, doi:10.1080/02626667.2010.504677.

Kundzewicz, Z. W. (2002), Floods in the context of climate change and variability, in *Climatic Change: Implications for the Hydrological Cycle and for Water Management*, edited by M. Beniston, pp. 225–247, Kluwer Academic Publishers, Netherlands.

Kysely, J. (2008), A Cautionary Note on the Use of Nonparametric Bootstrap for Estimating Uncertainties in Extreme-Value Models, *Journal of applied meteorology and climatology*, *47*, 3236–3251, doi:10.1175/2008JAMC1763.1.

Lamb, R., and A. L. Kay (2004), Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain, *Water Resources Research*, *40*, 1–13, doi:10.1029/2003WR002428.

Lang, M., T. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling, *Journal of Hydrology*, *225*, 103–117.

López-Pintado, S., and J. Romo (2009), On the Concept of Depth for Functional Data, *Journal of the American Statistical Association*, *104*, 718–734, doi:10.1198/jasa.2009.0108.

Madsen, H., P. F. Rasmussen, and D. Rosbjerg (1997), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. At-site modeling, *Water Resources Research*, *33*(4), 747–757.

Marshall, L., D. Nott, and A. Sharma (2005), Hydrological model selection: A Bayesian alternative, *Water Resources Research*, *41*(10), 1–11, doi:10.1029/2004WR003719.

Martins, E. S., and J. R. Stedinger (2001a), Generalized Maximum Likelihood Pareto-Poisson estimators for partial duration series, *Water Resources Research*, *37*(10), 2551–2557, doi:10.1029/2001WR000367.

Martins, E. S., and J. R. Stedinger (2001b), Historical information in a generalized Maximum Likelihood Framework with partial duration and annual maximum series, *Water Resources Research*, *37*(10), 2559–2567, doi:10.1029/2000WR000009.

McMillan, H. K., and I. K. Westerberg (2015), Rating curve estimation under epistemic uncertainty, *Hydrological Processes*, *29*(7), 1873–1882, doi:10.1002/hyp.10419.

Merz, B., and A. H. Thieken (2005), Separating natural and epistemic uncertainty in flood frequency analysis, *Journal of Hydrology*, *309*(1-4), 114–132, doi:10.1016/j.jhydrol.2004.11.015.

Merz, B., H. Kreibich, and H. Apel (2008), Flood risk analysis: uncertainties and validation, *Österreichische Wasser- und Abfallwirtschaft*, *60*(5-6), 89–94, doi:10.1007/s00506-008-0001-4.

Merz, R., and G. Blöschl (2003), A process typology of regional floods, *Water Resources Research*, *39*(12), 1340, doi:10.1029/2002WR001952.

Meylan, P., A.-C. Favre, and A. Musy (2012), *Predictive hydrology. A frequency analysis approach*, 212 pp., Science Publishers. St. Helier, Jersey, British Channel Islands.

Montanari, A., and D. Koutsoyiannis (2012), A blueprint for process-based modeling of uncertain hydrological systems, *Water Resources Research*, *48*(9), 1–15, doi:10.1029/2011WR011412.

Nadarajah, S. (2007), Probability models for unit hydrograph derivation, *Journal of Hydrology*, *344*, 185–189, doi:10.1016/j.jhydrol.2007.07.004.

Nearing, G. S., Y. Tian, H. V. Gupta, M. P. Clark, K. W. Harrison, and S. V. Weijs (2016), A Philosophical Basis for Hydrologic Uncertainty, *Hydrological Sciences Journal*, *6667*(May), 1–20, doi:10.1080/02626667.2016.1183009.

Pappenberger, F., and K. J. Beven (2006), Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resources Research*, *42*(5), 1–8, doi:10.1029/2005WR004820.

Pilgrim, D. H. (1986), Bridging the Gap Between Flood Research and Design Practice, *Water Resources Research*, *22*(9), 165–176.

Poulin, A., D. Huard, A.-C. Favre, and S. Pugin (2007), Importance of tail dependence in bivariate frequency analysis, *Journal of Hydrologic Engineering*, *12*(4), doi:10.1061/(ASCE)1084-0699(2007)12:4(394).

Qi, W., C. Zhang, G. Fu, and H. Zhou (2016), Imprecise probabilistic estimation of design floods with epistemic uncertainties, *Water Resources Research*, *52*, 4823–4844, doi:10.1002/2015WR017663.

Rai, R. K., S. Sarkar, and V. P. Singh (2009), Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph, *Water Resources Management*, *23*, 899–929, doi:10.1007/s11269-008-9306-0.

Requena, A. I., L. Mediero, and L. Garrote (2013), A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation, *Hydrological Earth System Sciences*, *17*, 3023–3038, doi:10.5194/hess-17-3023-2013.

Schumann, A. H., D. Nijssen, and M. Pahlow (2010), Handling uncertainties of hydrological loads in flood retention planning, *International Journal of River Basin Management*, *8*(3-4), 281–294, doi:10.1080/15715124.2010.512561.

Serinaldi, F. (2009), Assessing the applicability of fractional order statistics for computing confidence intervals for extreme quantiles, *Journal of Hydrology*, *376*, 528–541, doi:10.1016/j.jhydrol.2009.07.065.

Serinaldi, F. (2013), An uncertain journey around the tails of multivariate hydrological distributions, *Water Resources Research*, *49*(10), 6527–6547, doi:10.1002/wrcr.20531.

Serinaldi, F. (2015), Dismissing return periods!, *Stochastic Environmental Research and Risk Assessment*, *29*, 1179–1189, doi:10.1007/s00477-014-0916-1.

Shu, C., and T. Ouarda (2008), Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system, *Journal of Hydrology*, *349*, 31–43, doi:10.1016/j.jhydrol.2007.10.050.

Sikorska, A., and B. Renard (2017), Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges, *Advances in Water Resources*, *105*, 51–66, doi:10.1016/j.advwatres.2017.04.011.

Sikorska, A. E., A. Scheidegger, K. Banasik, and J. Rieckermann (2012), Bayesian uncertainty assessment of flood predictions in ungauged urban basins for conceptual rainfall-runoff models, *Hydrology and Earth System Sciences*, *16*(4), 1221–1236, doi:10.5194/hess-16-1221-2012.

Sikorska, A. E., A. Scheidegger, K. Banasik, and J. Rieckermann (2013), Considering rating curve uncertainty in water level predictions, *Hydrology and Earth System Sciences*, *17*, 4415–4427, doi:10.5194/hess-17-4415-2013.

Sikorska, A. E., D. Viviroli, and J. Seibert (2015), Flood type classification in mountainous catchments using crisp and fuzzy decision trees, *Water Resources Research*, *51*(10), 7959–7976, doi:10.1002/2015WR017326.

Singh, V. P., S. X. Wang, and L. Zhang (2005), Frequency analysis of nonidentically distributed hydrologic flood data, *Journal of Hydrology*, *307*(1-4), 175–195, doi:10.1016/j.jhydrol.2004.10.029.

Strupczewski, W., V. Singh, and S. Węglarczyk (2002), Asymptotic bias of estimation methods caused by the assumption of false probability distribution, *Journal of Hydrology*, *258*, 122–148.

Sun, H., T. Jiang, C. Jing, B. Su, and G. Wang (2017), Uncertainty analysis of hydrological return period estimation, taking the upper Yangtze River as an example, *Hydrology and Earth System Sciences Discussions*, pp. 1–26, doi:10.5194/hess-2016-566.

Tisseuil, C., M. Vrac, S. Lek, and A. J. Wade (2010), Statistical downscaling of river flows, *Journal of Hydrology*, *385*(1-4), 279–291, doi:10.1016/j.jhydrol.2010.02.030.

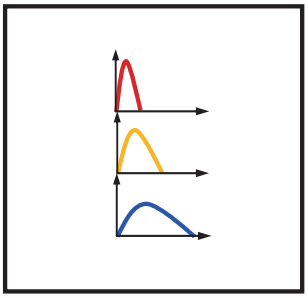
Tung, Y.-K., and B.-C. Yen (2005), *Hydrosystems engineering uncertainty analysis*, 285 pp., McGraw-Hill Book Company, New York.

Wasserman, L. (2006), *All of nonparametric statistics*, 269 pp., Springer, New York.

Xu, Y.-P., M. J. Booij, and Y.-B. Tong (2010), Uncertainty analysis in statistical modeling of extreme hydrological events, *Stochastic Environmental Research and Risk Assessment*, *24*, 567–578, doi:10.1007/s00477-009-0337-8.

Yue, S., T. Ouarda, B. Bobée, P. Legendre, and P. Bruneau (2002), Approach for describing statistical properties of flood hydrograph, *Journal of Hydrologic Engineering*, *7*(2), 147–153, doi:10.1061/(ASCE)1084-0699(2002)7:2(147).

# Paper V



# Identification of flood reactivity regions via the functional clustering of hydrographs

Manuela I. Brunner[1,2], Daniel Viviroli[1,3], Reinhard Furrer[4], Jan Seibert[1,5], and Anne-Catherine Favre[2]

[1]Department of Geography, University of Zurich, Zurich, Switzerland
[2]Univ. Grenoble-Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[3]belop gmbh, Sarnen, Switzerland
[4]Department of Mathematics, University of Zurich, Zurich, Switzerland
[5]Department of Earth Sciences, Uppsala University, Uppsala, Sweden

*Water Resources Research, accepted*

**Abstract**

Flood hydrograph shapes contain valuable information on the flood-generation mechanisms of a catchment. To make good use of this information, we express flood hydrograph shapes as continuous functions using a functional data approach. We propose a clustering approach based on functional data for flood hydrograph shapes to identify a set of representative hydrograph shapes on a catchment scale and use these catchment-specific sets of representative hydrographs to establish regions of catchments with similar flood reactivity on a regional scale. We applied this approach to flood samples of 163 medium-size Swiss catchments. The results indicate that three representative hydrograph shapes sufficiently describe the hydrograph shape variability within a catchment and therefore can be used as a proxy for the flood behavior of a catchment. These catchment-specific sets of three hydrographs were used to group the catchments into three reactivity regions of similar flood behavior. These regions were not only characterized by similar hydrograph shapes and reactivity but also by event magnitudes and triggering event conditions. We envision these regions to be useful in regionalization studies, regional flood frequency analyses, and to allow for the construction of synthetic design hydrographs in ungauged catchments. The clustering approach based on functional data which establishes these regions is very flexible and has the potential to be extended to other geographical regions or towards the use in climate impact studies.

# 1 Introduction

Hydrological processes vary widely from one environment to the next (*McDonnell and Woods*, 2004), which causes distinct flood responses (*Merz and Blöschl*, 2009). It is often useful to reduce this variability by grouping catchments with similar governing processes. The establishment of regions which are similar in terms of their flood behavior is challenging but is an important first step in regionalization studies (*Prinzio et al.*, 2011; *Blöschl et al.*, 2013) and regional flood frequency analyses that allow for the estimation of flood frequencies in ungauged catchments (*Hosking and Wallis*, 1997). The often-used index flood method (*Dalrymple*, 1960), for instance, is based on regions consisting of hydrologically similar catchments. It uses information from sites within a given region to estimate the magnitude of extreme events corresponding to a predefined return period at a target site (*Requena et al.*, 2017). Similarity measures typically used to delineate regions of similar hydrological behavior are physiographical (e.g. catchment area or altitude), climatological (e.g. daily rainfall statistics), and/or hydrological (e.g. mean daily flow) characteristics (see *Ali et al.* (2012) for an overview on variables used in previous studies). Preferably, such regions are established using physiographical or climatological catchment characteristics since these are also available for ungauged catchments (*Acreman and Sinclair*, 1986; *Hosking and Wallis*, 1997; *Ilorme and Griffis*, 2013), which enables the attribution of an ungauged catchments to an existing region. However, physiographical and climatological catchment similarity do not often correspond to hydrological similarity (*Oudin et al.*, 2010; *Ali et al.*, 2012) and even less to similarity in the flood behavior of a catchment (*Merz and Blöschl*, 2009). Therefore, hydrologically similar regions are often delineated based on hydrological catchment characteristics or runoff signatures (*Burn and Boorman*, 1992), such as the median daily flow, annual runoff coefficient, slope of the flow duration curve (*Boscarello et al.*, 2016), seasonality indices (*Castellarin et al.*, 2001), monthly Pardé coefficients (*Hailegeorgis and Alfredsen*, 2017), or statistical measures which reflect the shape of the flood distribution (*Hosking and Wallis*, 1997). The latter characteristics usually focus on flood peaks and reflect only a part of the flood behavior of a catchment. The focus on flood peaks neglects other hydrograph characteristics such as volume and shape, which are equally important for many flood risk management tasks, especially for those involving storage (*Pilgrim*, 1986; *Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall*, 2012), and potentially provide crucial information on the flood behavior of a catchment. These hydrograph characteristics could be useful for the identification of regions similar in terms of their flood behavior since the hydrograph integrates temporal and spatial variations in water input, storage, and water processes within a catchment (*Hannah et al.*, 2000). The goal of this study is to employ the information integrated in hydrograph shapes for the identification of regions with a similar flood behavior. We propose a catchment clustering scheme consisting of two steps that accounts for different flood mechanisms acting within a catchment. Clustering is often used as an exploratory tool to identify distinct groups so that the observations within each group are similar to each other while observations in different groups are different from each other (*James et al.*, 2013). The clusters resulting from applying a cluster algorithm can not be validated directly but indirectly by their interpretability and usefulness (*Webster and Oliver*, 2007). The first step in the clustering scheme is the identification of *representative* hydrograph shapes using observed flood event hydrograph shapes. The term representative is used for shapes that describe the hydrograph shape variability and potential flood responses or causative mechanisms within a catchment. Compared to existing flood process classification schemes, such as the one proposed by *Merz and Blöschl* (2003), no meteorological information in addition to the streamflow data is used to identify different flood event types. The second step in the clustering scheme identifies regions of catchments with a similar flood reactivity by clustering the catchment-specific sets of representative hydrograph shapes obtained in the first step.

The approach is based on the clustering of hydrograph shapes represented as functional data (FD). In contrast to classical multivariate data, FD is continuously defined and does not depend on the choice of several hydrograph characteristics, such as peak discharge, hydrograph volume, duration, or a few parameters representing the shape of a hydrograph (*Yue et al.*, 2002), but instead uses the whole information stored in the hydrograph (*Chebana et al.*, 2012). FD analysis is more general, flexible, and representative of the real hydrological phenomena than classical multidimensional analysis and avoids the subjective choice of a set of hydrograph characteristics (*Ternynck et al.*, 2016). FD are conceptually defined in a continuous framework. In practice however, they are usually observed at discrete points in time and stored in a finite-dimensional way. Hydrographs can be considered as FD since they fulfill this criterion. The first step in FD analysis is often the reconstruction of the

functional form of data from discrete observations (see Figure 1 for an illustration). Most commonly, this is done by considering the data as part of a finite dimensional space spanned by some basis functions. Alternatively, the data can be smoothed non-parametrically (*Jacques and Preda*, 2014). This allows the representation of an individual functional datum as a continuous function rather than as its values at particular points (*Ramsay and Silverman*, 2002). It has been shown in previous

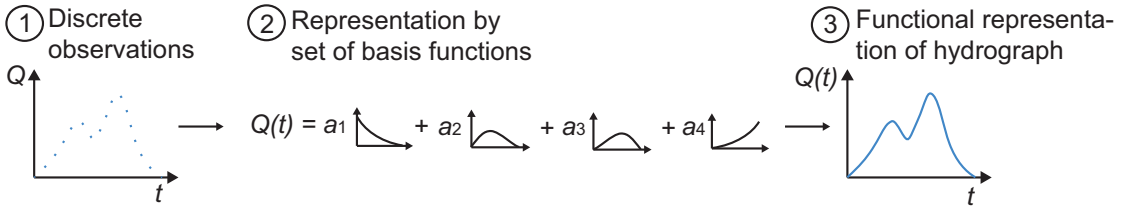



Figure 1: Getting from discrete measurements (1) to a functional representation of a hydrograph (3) by representing the data by a set of basis functions (2).

studies that the FD framework can be beneficial in the identification of groups of similar hydrographs over a range of temporal scales, such as yearly hydrographs (*Merleau et al.*, 2007; *Jamaludin*, 2016), spring flood events (duration of six months) (*Ternynck et al.*, 2016), and diurnal discharges (duration of one day) (*Hannah et al.*, 2000).

In this study, we adapt the FD framework to cluster flood event hydrograph shapes (duration of three days) to identify catchment specific sets of hydrograph shapes and use these sets to establish homogeneous regions in terms of flood reactivity. Well-known clustering algorithms such as hierarchical and $k$-means algorithms can be adapted to the case of FD (*Cuevas*, 2014). *Jacques and Preda* (2014) grouped the major approaches into four categories: 1) raw data clustering, 2) two-stage approaches which first reduce the dimension of the data and second perform clustering, 3) nonparametric clustering approaches, and 4) model-based clustering approaches. An initial analysis, where each of these method categories was tested, showed that two-stage approaches were most suitable to cluster hydrographs since they resulted in meaningful clusters. We therefore focused on this type of methods.

# 2 Data

## 2.1 Study catchments

This study was performed using runoff data from 163 Swiss catchments (see Figure 2) with a wide range of catchment characteristics and flood behaviors. The selected catchments have hourly flow series of at least 20 years in duration and ranging up to 53 years. The catchments' runoff is neither significantly altered by regulated lakes upstream or inland canals nor by urbanized areas or hydropower. The catchments are small to medium-size (6 to 1800 km$^2$), situated between 400 and 2600 m.a.s.l. (mean elevation), and either have no or only small areas with glaciers. The catchment set covers a wide range of runoff regimes and catchment characteristics and is therefore well suited for illustrating the proposed approach.

## 2.2 Flood events

The basis for the functional hydrograph analysis was samples of flood events extracted from the runoff time series of the 163 study catchments. To sample flood events, we used a peak-over-threshold approach based on the procedure proposed by *Lang et al.* (1999). The threshold for the peak discharge was chosen iteratively to fulfill a target condition of four events per year on average which is a trade-off between maximizing the information content in the sample and keeping the assumption of independence between events. For each of these events, sampled according to the flood peaks, the flood volume and hydrograph were determined over a fixed event window of 72 h. Flood events that ended in a secondary peak were removed automatically from the dataset to ensure that events triggered by two independent convective precipitation events were not considered as one event. The baseflow was separated from the direct flow using a recursive digital filter (*Eckhardt*, 2005). The

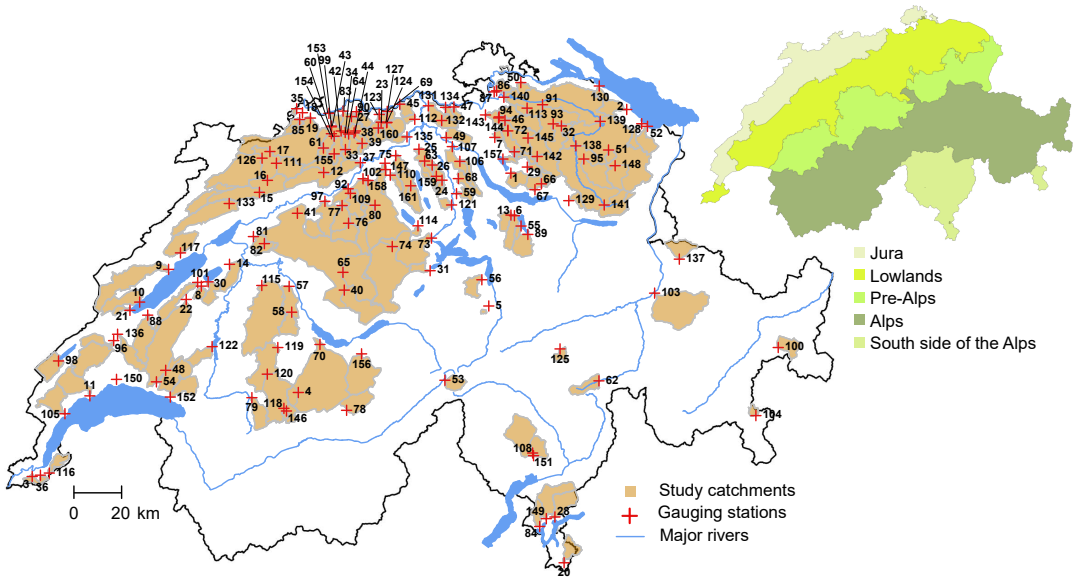



Figure 2: 163 study catchments in Switzerland and geographical regions of Switzerland as introduced by *Bundesamt für Umwelt BAFU and Eidg. Forschungsanstalt WSL* (2012).

resulting direct flow component of the hydrographs was then normalized so that the volume of the modified hydrographs was equal to one. This was done by dividing the ordinate of each hydrograph by the volume $V$. In the remainder of this paper, we refer to these normalized hydrographs as *hydrograph shapes.*The advantage of working with normalized hydrographs is that hydrograph shapes of catchments with different sizes can directly be compared. We refer the reader to *Brunner et al.* (2017) for a more detailed description of the flood sampling and baseflow separation procedures.

# 3 Methods

In this study, we propose a catchment clustering scheme for the establishment of regions with similar flood behaviors (for an illustration see Figure 3). The framework consists of two main parts:

1. The identification of sets of representative hydrograph shapes within individual catchments via a clustering approach in which hydrograph shapes are represented as FD. This first step allows for the characterization of the flood behavior of a catchment, that is typically influenced by several causative mechanisms (*Merz and Blöschl*, 2003), via several representative hydrograph shapes.

2. The establishment of regions with similar flood behaviors by clustering the catchment-specific sets of representative hydrograph shapes obtained in Step 1.

The methodology was established on the set of 163 Swiss catchments introduced above, however, its general principle is transferable to other geographical regions. The color coding used in this paper clearly distinguishes the clusters resulting from Steps 1 and 2. Results referring to the clusters established in Step 1 are displayed using saturated colors (red, orange, and blue) while those referring to the clusters established in Step 2 are displayed in pastel colors (rose, yellow, and light blue) (see Figure 3).

## 3.1 Identification of representative hydrograph shapes

The first part of this analysis focused on the identification of representative hydrograph shapes at a catchment scale. We used a clustering approach as an exploratory tool to identify the number of hydrograph shapes necessary to sufficiently describe the variability of hydrograph shapes within each of the study catchments. Contrary to classical clustering approaches, hydrograph shapes were not considered as multivariate data but expressed as FD, i.e., each hydrograph was represented as a

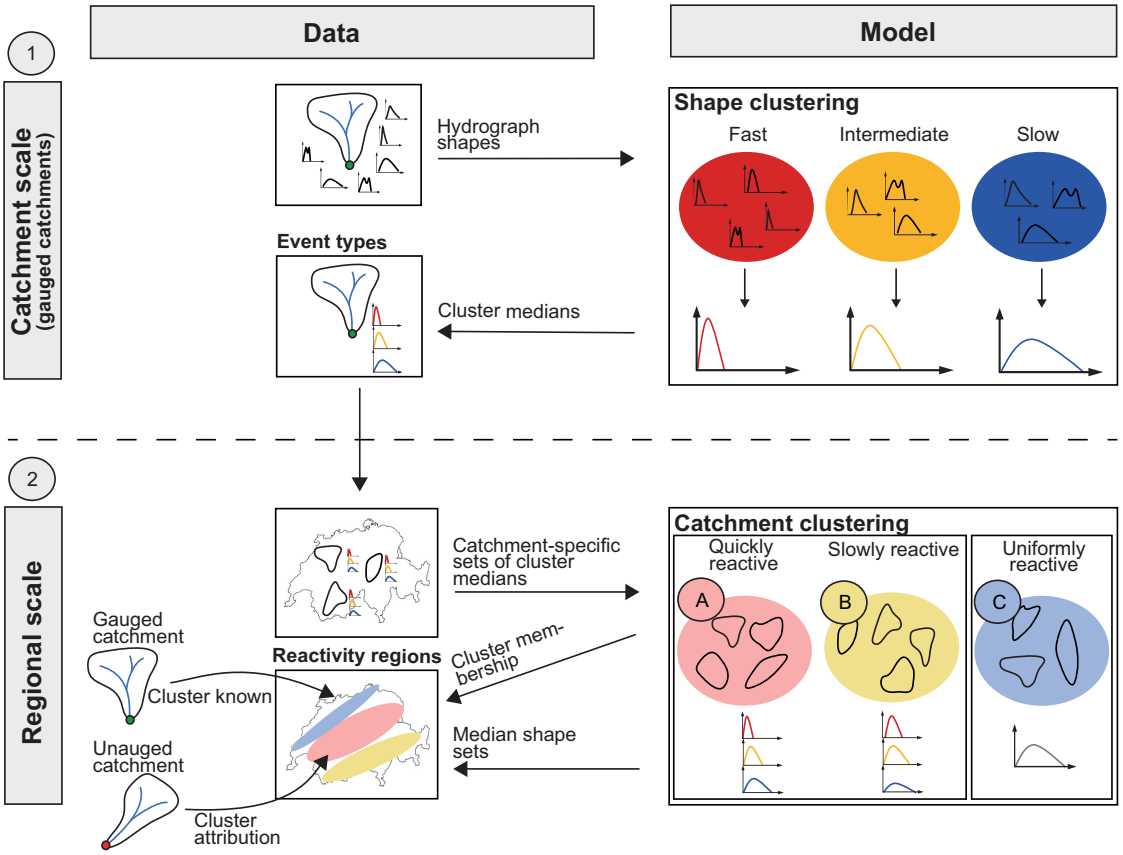



Figure 3: Illustration of the clustering framework. The data input and output for the models are indicated for the two different parts: 1) hydrograph shape clustering, 2) catchment clustering. Hydrograph shape clustering is done on the catchment scale and catchment clustering on the regional scale.

function of time. The normalized hydrographs were projected on a set of basis functions to represent them as FD in a simple and low-dimensional framework. A basis function system is a set of known orthogonal functions. Any function can be approximated arbitrarily well by a weighted sum or by a linear combination of a sufficiently large number of functions out of such a set of basis functions. Spline functions are the most common choice of basis functions for non-periodic FD. We used a set of B-spline functions (Data2fd from fda package in *R* (*Ramsay et al.*, 2014)) as basis functions because they are able to mimic the main characteristics of hydrograph shapes (*Abraham et al.*, 2003). A (smoothing) spline function is determined by the order of the polynomial segments and the number and placement of knots. The number of knots determines the ability of spline functions to represent sharp features in a curve and the knots can be placed such that they are denser in areas with stronger variations than in smooth areas (*Höllig and Hörner*, 2013). Increasing the number of splines does not always improve the fit to the data because the functional space defined by $n$ basis or B-splines is not necessarily contained within the one defined by $(n+1)$ B-splines (*Ramsay and Silverman*, 2005). We chose a set of four B-splines of order four (Figure 4). B-splines of order four were sufficiently flexible to represent the hydrograph shapes under study. We did not go beyond this number to avoid unnecessary complexity. Similarly, an increase of the number of basis functions above four did not further improve the representation of hydrograph shapes as FD. We then computed the coefficients for each of the normalized hydrographs for the four B-Spline bases.

The set of four coefficients per hydrograph shape was used as an input for the cluster analysis. The coefficient sets were clustered using the classical $k$-means algorithm. We applied the algorithm for $k = 2$ up to $k = 6$ number of clusters. The resulting clusters were assessed graphically for their suitability to represent the main part of the hydrograph shape variability within a catchment. In addition to this graphical assessment, we computed the silhouette widths over all catchments and hydrograph shapes (*Rousseeuw*, 1987) for different numbers of clusters. The silhouette width shows

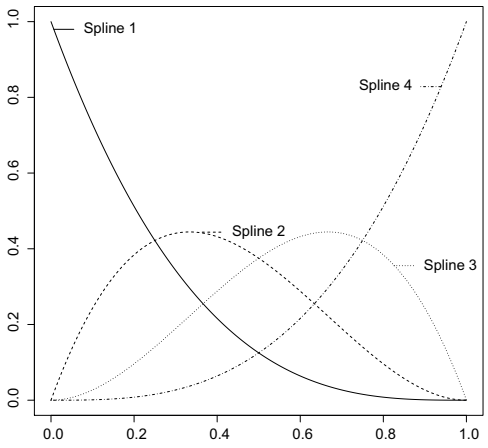



Figure 4: Four B-splines of order four used for the representation of hydrograph shapes as FD.

which hydrograph shapes lie well within their clusters and which ones lie in between clusters. The overall silhouette width provides an evaluation of clustering validity. Three clusters were found to represent hydrograph shape variability well and increasing the number of clusters did not further improve this representation (i.e., clusters with very few hydrograph shapes were built) and led to a decrease in the overall silhouette width. We decided to fix the number of clusters to three for all catchments since this allowed for their comparison across different catchments. Three clusters were a compromise between over representing hydrograph shape variability in catchments with a rather low variability and under representing variability in catchments with a rather high shape variability. Each of the three clusters consisted of a set of similar hydrograph shapes, which was summarized by their median hydrograph. We used the $h$-mode depth to order the hydrographs in the sample and to identify the median hydrograph within the set of similar hydrographs (*Cuevas et al.*, 2007). The concept of data depth aims at measuring the centrality of a given curve (in our case, the hydrographs) within a group of curves and can be used to define the ranks of functional data (*Fraiman and Muniz*, 2001) and therewith robust estimators of a location parameter such as the median or the trimmed mean. The median hydrograph identified using the $h$-mode depth is actually an observed hydrograph, which would not be the case if the median hydrograph was defined by the median flow observation at each time step. The three resulting median hydrographs (one for each cluster) were said to be representative of the hydrograph shapes within the catchment. They were ranked according to their increasing time to peak to order them according to their reaction time, and were called fast, intermediate, and slow shape, respectively. The set of these three representative hydrograph shapes was used, in the following step, to describe the general flood behavior of a catchment. This set represents the variability in observed hydrograph shapes and, in the remainder of this paper, we will refer to the three representative hydrograph shapes composing it as the fast, intermediate, and slow *event types*.

## 3.2 Establishment of regions with similar flood behaviors

The second part of this analysis focused on the regional scale to identify regions of catchments with a similar flood behavior in terms of their three representative hydrograph shapes (fast, intermediate, and slow). While these representative hydrograph shapes are quite distinct in most catchments, there are catchments where they are quite similar and a distinction between three shapes is not appropriate. The identification of regions with a similar flood behavior therefore consisted of two steps: 1) the identification of a region of catchments with a uniform flood reaction and 2) the identification of regions with similar flood reactions among the remaining catchments. The second step focused on clustering the catchments according to their three representative hydrograph shapes.

1. **Identification of a region with a uniform flood reaction:** We first identified catchments where the three representative hydrograph shapes were similar. These catchments were said to exhibit a uniform flood reaction and together built the region of uniformly reactive catchments. To identify the uniformly reactive catchments, we looked at the time to peak of the median

hydrographs of the clusters obtained in the first part of the analysis. We computed the sum of the differences between the time to peak of the three representative hydrographs. The sum of the differences was found to be small in the catchments where the three hydrograph shapes were not very distinct. The threshold for the sum of the differences separating the uniformly reactive from the remaining catchments was set to a threshold of five based on visual inspection i.e., an assessment of the similarity of three representative hydrograph shapes and the sum of differences between their times to peak. The threshold was set such that catchments where all three shapes were similar were said to be uniformly reactive and those where shapes started to differ were said to be non-uniformly reactive. The uniformly reactive catchments built their own region and were characterized by only one representative hydrograph shape.

2. **Identification of regions with similar flood reactions:** In most catchments, the hydrograph shape variability could only be well represented by three hydrograph shapes: a fast, an intermediate, and a slow hydrograph. However, a fast hydrograph in one catchment was not necessarily a fast hydrograph in another catchment. Contrarily, catchments showed differences in their general runoff reaction time. To identify regions of catchments with a similar reactivity, we applied clustering to their sets of representative hydrograph shapes. Catchments with similar sets of representative hydrograph shapes were said to be of similar reactivity and similar climate conditions. Working with normalized shapes allowed for the identification of catchments with a similar reactivity independent of their catchment size. For the clustering, the three representative hydrographs per catchment were again expressed as FD and we again computed coefficients for four B-splines of order four. Similar to Step 1, neither increasing the number of spline bases nor their order further improved the clustering results, which was also confirmed by the overall silhouette width. The four coefficients for each of the representative hydrograph shapes were stacked together to form a vector of $4 \times 3 = 12$ coefficients. We applied hierarchical clustering instead of $k$-means clustering to allow for non-elliptical clusters (*Gordon*, 1999). As an objective function, we used Ward's minimum variance criterion, which minimizes the total within-cluster variance (*Ward*, 1963). The hierarchical clustering tree was symmetrical which suggested cuts at either $k = 2$ or $k = 4$. We found that two clusters were sufficient to distinguish catchments with a generally fast reaction (quickly responding fast, intermediate, and slow events) from catchments with a generally slow reaction (delayed fast, intermediate, and slow events). This was confirmed by the overall silhouette width which was highest for $k = 2$ and decreased with an increase of the number of clusters.

This second part of the analysis divided the study domain into three regions with a similar reactivity, which we herein refer to as *reactivity regions*. A region where catchments showed a fast reaction to rainfall events, a region where catchments showed a slow reaction to rainfall events, and a region where catchments were characterized by a relatively uniform reaction to precipitation events.

## 3.3 Event condition analysis

The hydrograph clustering approaches proposed in this study first divided the events observed in a catchment into three event types and second, divided the catchments into three regions of similar reactivity based on hydrograph shapes, without considering event magnitudes or pre-event conditions. A direct validation of the results is neither possible for the event type clusters nor the reactivity regions since a cluster analysis is exploratory (*James et al.*, 2013) and the "true" cluster memberships are not known. The validity of clusters was therefore indirectly assessed by their suitability to form meaningful clusters in relation to hydro-meteorological conditions. Event types might be distinct in their triggering precipitation or antecedent wetness conditions. We looked at the precipitation events triggering the individual flood events and at their antecedent wetness conditions to investigate the link between event type, region, and event conditions. Event precipitation and antecedent wetness were computed using hourly gridded precipitation data. We used the grid-data product CombiPrecip provided by *MeteoSwiss* (2013), which was computed using a geostatistical approach combining rain-gauge measurements and radar estimates and is available from the year 2005 at a spatial resolution of 1 km. Continuous precipitation time series from 2005 to 2014 were obtained for each catchment by simply averaging the precipitation of the individual grid-cells lying within the catchment. These time series were used to compute the current precipitation index ($I_{\mathrm{CP}}$), which reflects the current catchment wetness. It is defined as a continuous function of precipitation, which accumulates on rainy days and

exponentially decays during the periods of no rainfall with a recession coefficient of 0.9 (*Smakhtin and Masse*, 2000). We then defined the antecedent wetness condition prior to a flood event as the $I_{CP}$ at the beginning of the event. The time series were also used to compute two characteristics related to the event triggering precipitation: total precipitation amount and maximum hourly precipitation intensity. The precipitation data related to the sampled flood events was defined over a window of a maximum of 84 hours starting 12 hours before the onset of the flood event and ending with the flood event. The three characteristics, antecedent wetness, total event precipitation, and maximum hourly event precipitation, were used to analyze differences in the triggering mechanisms of flood events belonging to the three different event types occurring in the three regions with similar flood behavior.

# 4 Results

## 4.1 Identification of representative hydrograph shapes

The identification of clusters of representative hydrograph shapes within a catchment was based on a functional representation of the hydrograph shapes with respect to four B-Spline basis functions. The coefficients of these B-splines allowed for the separation of fast, intermediate, and slow events within a catchment (see Figure 5 showing the Langete-Huttwil as an example catchment). The fast events were characterized by rather steep rising and falling limbs of the hydrographs. Intermediate events typically also showed quite steep rising limbs while the falling limbs were less steep than the ones observed for the fast events. Slow events were typically characterized by both elongated rising and falling limbs. The fast events were clearly related to a season of occurrence, namely, summer (red curves in the upper panel of Figure 5) while intermediate and slow events occurred throughout the seasons. The clusters established via the hydrograph shape also had a meaning in terms of event magnitude even though they have been established based on the normalized hydrograph shape information only (see the boxplots of peak discharges and hydrograph volumes in the lower panel of Figure 5). Events in the fast event cluster were generally characterized by high peak discharges but low hydrograph volumes while events in the slow event cluster showed high volumes but low peak discharges. The magnitudes of intermediate events lay in between those of the fast and slow events. This general pattern was visible in most catchments of the dataset even though it was not always as clear as in the case of the Langete-Huttwil catchment (Figure 5).

The three median hydrographs of the fast, intermediate, and slow event clusters together built the catchment-specific set of representative hydrograph shapes (see Figure 6 for three examples). While this set consisted of three distinct hydrograph shapes in most catchments, one hydrograph was per definition (see Section 3.2) sufficient to describe the runoff reaction in the uniformly reactive catchments (see Figure 6C representing the Birs-Münchenstein). On the contrary, the remaining catchments showed different types of runoff reactions. However, a fast hydrograph in one catchment (e.g. Figure 6B representing Arbogne-Avenches) would rather represent a slow hydrograph in another catchment (e.g. Figure 6A illustrating Alp-Einsiedeln). We concluded that catchments can be generally characterized by different event-reaction times since fast, intermediate, and slow hydrographs looked differently in different catchments.

## 4.2 Establishment of regions with similar flood behaviors

The catchment-specific sets of representative hydrograph shapes were used to establish regions with similar flood behaviors. The 11 catchments with a set of very similar hydrograph shapes were said to belong to a uniformly reactive cluster (see Figure 7C). The remaining catchments were then used to establish regions of catchments with similar flood behaviors. The clustering of the catchment-specific event sets resulted in two clusters with distinct flood-event reaction-times: 71 quickly reactive catchments (see Figure 7A) and 81 slowly reactive catchments (see Figure 7B). The quickly reactive catchments were mainly located in the Swiss Plateau, slowly reactive catchments in the Jura mountains, the Alpine region, and the Swiss Plateau, and uniformly reactive catchments mainly in the Jura mountains.

At a regional level, catchment-specific sets of representative hydrographs were distinct between catchments of a generally quick and a generally slow flood runoff reaction. The B-spline coefficients of these catchment-specific sets indicated which hydrograph features were important in distinguishing

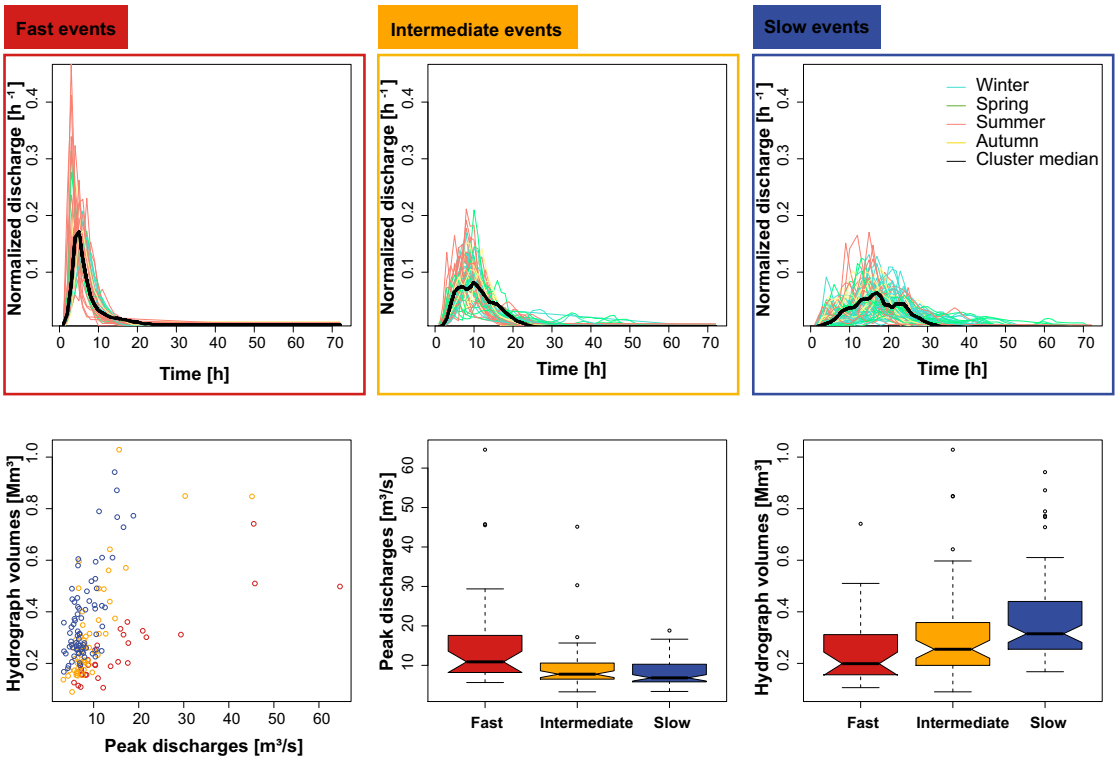



Figure 5: Three clusters of hydrograph shapes for the Langete-Huttwil catchment. The events belonging to the fast, intermediate, and slow cluster are displayed according to season using four colors (top panel). The median hydrograph shape per cluster is indicated in black. The characteristics of the events belonging to the three clusters in terms of the magnitudes of peak discharges and hydrograph volumes are displayed in a scatterplot and as boxplots (lower panel). The notches of the boxplots indicate the confidence intervals around the medians.

quickly reactive from slowly reactive catchments. Figure 8 shows the coefficients of the four B-splines for the representative hydrographs of the catchments in the quickly and slowly reactive region for the three event type classes fast, intermediate, and slow. The coefficients for Splines 1 to 4 (see Figure 4) of the representative fast, intermediate, and slow hydrographs were different for catchments in the quickly reactive and catchments in the slowly reactive region (Figure 8). These differences were the most expressed for the fast events and less expressed for the intermediate and slow events. In summary, quickly reactive catchments were characterized by hydrograph shapes that can be described as a sum of Splines 1 and 3 with a high coefficient and Splines 2 and 4 with a low coefficient. The opposite was the case for hydrographs found in catchments belonging to the slowly reactive region.

The shape clusters and their medians were representative for their region (Figure 9). The catchments in the quickly reactive and the slowly reactive regions (see Figure 7) could be characterized by a set of representative hydrograph shapes while the catchments in the uniformly reactive region were sufficiently described by one hydrograph shape. Catchments in the quickly reactive region were characterized by hydrograph sets with a steeper recession limb than the catchments in the slowly reactive region. Figure 10 shows that the reactivity regions not only differed in terms of their representative hydrograph shapes but also in terms of hydrograph magnitudes. Events occurring in the quickly reactive catchments were characterized by rather high peak discharges compared to flood volumes while events occurring in the slowly reactive catchments showed rather high volumes compared to peak discharges.

### 4.2.1 Event conditions

The events occurring in the three regions and belonging to the three event types fast, intermediate, and slow not only showed different runoff characteristics (magnitudes and shapes) but differed in their pre-event conditions (see Section 3.3) in the form of their triggering precipitation events and

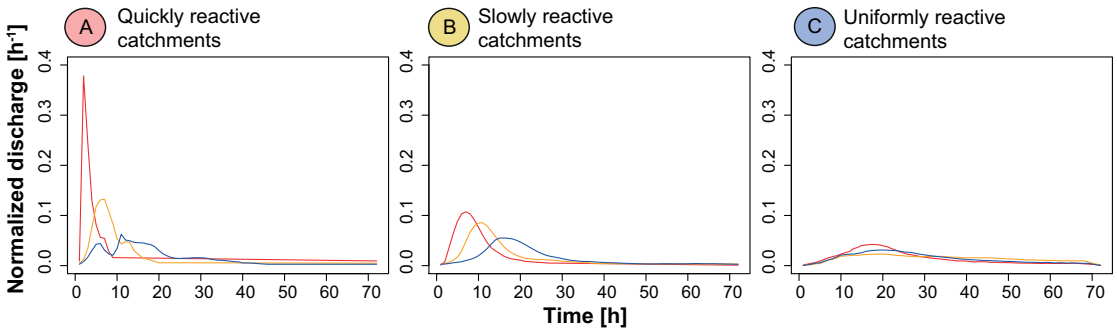



Figure 6: Fast, intermediate, and slow median hydrograph shapes for three catchments: A) Alp-Einsiedeln (quickly reactive region), B) Arbogne-Avenches (slowly reactive region), and C) Birs-Münchenstein (uniformly reactive region).

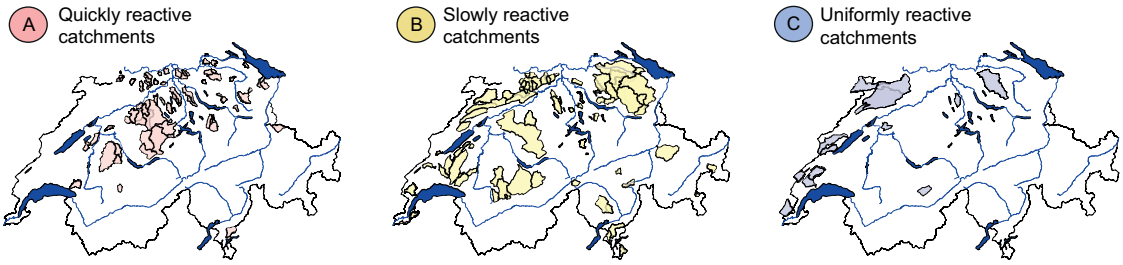



Figure 7: Catchments belonging to the quickly reactive (A), slowly reactive (B), and uniformly reactive (C) regions.

much less expressed in antecedent wetness (Figure 11). The flood events occurring in the three regions clearly differed in terms of their triggering precipitation. Flood events occurring in the quickly reactive region were characterized by higher maximum hourly precipitation intensities than catchments in the slowly and uniformly reactive regions. Within the regions, fast events showed clearly higher hourly precipitation intensities than intermediate and slow events. The triggering precipitation events for the floods occurring in the three regions also differed in terms of their total precipitation amount. Precipitation events triggering floods in the quickly reactive region had generally higher precipitation amounts than precipitation events triggering floods in the slowly and uniformly reactive regions. Within one region, fast events showed slightly lower triggering precipitation amounts than intermediate and slow events. Note, that the differences of maximum hourly precipitation intensities between event types within the uniformly reactive region were much less distinct than the ones in the quickly and slowly reactive regions. This indicates that a distinction between the three event types does indeed not make sense in the uniformly reactive region, also from a meteorological point of view. The antecedent wetness conditions differed only slightly between regions and event types. The differences in triggering precipitation amounts and intensities highlights that a subdivision of floods into three event types is meaningful and that the three regions have a hydro-meteorological meaning.

# 5 Discussion

## 5.1 Identification of representative hydrograph shapes

The approach proposed for the clustering of hydrograph shapes using FD is simple to apply and avoids the reduction of flood hydrographs to a few hydrograph characteristics, as done when applying standard multivariate clustering techniques. The results presented above showed that the hydrograph clustering approach based on functional data allows for the identification of three representative hydrograph shape classes within individual catchments. These classes are not only distinct in their event shapes but also their event magnitudes and triggering mechanisms. The first class identified consists of fast events with steep rising and falling limbs caused by precipitation events with a high intensity but rather low total amounts. These events mainly occur in summer, which in previously

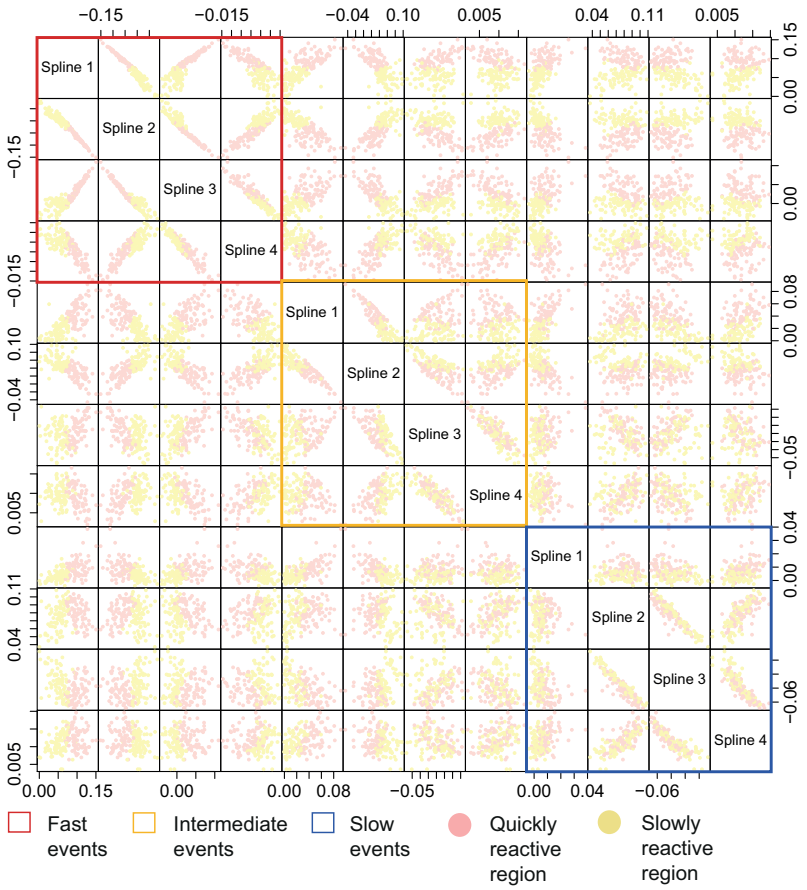



Figure 8: Coefficients for Splines 1 to 4 (see Figure 4) for the two reactivity regions quick and slow. The spline coefficients are given for fast, intermediate, and slow events separately. Scatterplots within the colored boxes (red, orange, and blue) indicate the relationship of the Splines 1 to 4 within one event type while the scatterplots outside the colored boxes indicate relationships between splines across different event types.

developed classification schemes correspond to flash floods (*Merz and Blöschl*, 2003; *Diezig and Weingartner*, 2007; *Sikorska et al.*, 2015). They are characterized by rather high peak discharges and low hydrograph volumes. The second event type class was composed of the intermediate events that were characterized by rather steep rising but elongated falling limbs and jointly high peak discharges and hydrograph volumes. Such events occur throughout the year and are triggered by precipitation events with medium intensity and amount. The third event class consisted of slow events with both elongated rising and falling limbs. These events typically showed high volumes but low peak discharges which is related to the precipitation events triggering them which are of low intensity but high total amount. The link between the triggering mechanism of the flood events and their shape and magnitude indicates that a hydrograph clustering approach based on functional data, as proposed in this study, is able to build event classes that are hydrologically and meteorologically meaningful. Eventually, the three event types could even be linked to hydro-meteorological patterns as suggested by *Nied et al.* (2014, 2017).

## 5.2 Establishment of regions with similar flood behaviors

Our results point out that catchment-specific sets of representative hydrograph shapes are useful in establishing regions of catchments with similar flood behaviors. Catchments with a set of three very similar hydrograph shapes, which could not be easily separated into fast, intermediate, and slow events, built the uniformly reactive region. The events occurring in the catchments belonging to this region were generally characterized by rather elongated rising and falling limbs, high volumes and low peak discharges. The catchments belonging to this region lie mainly in the Jura but also in

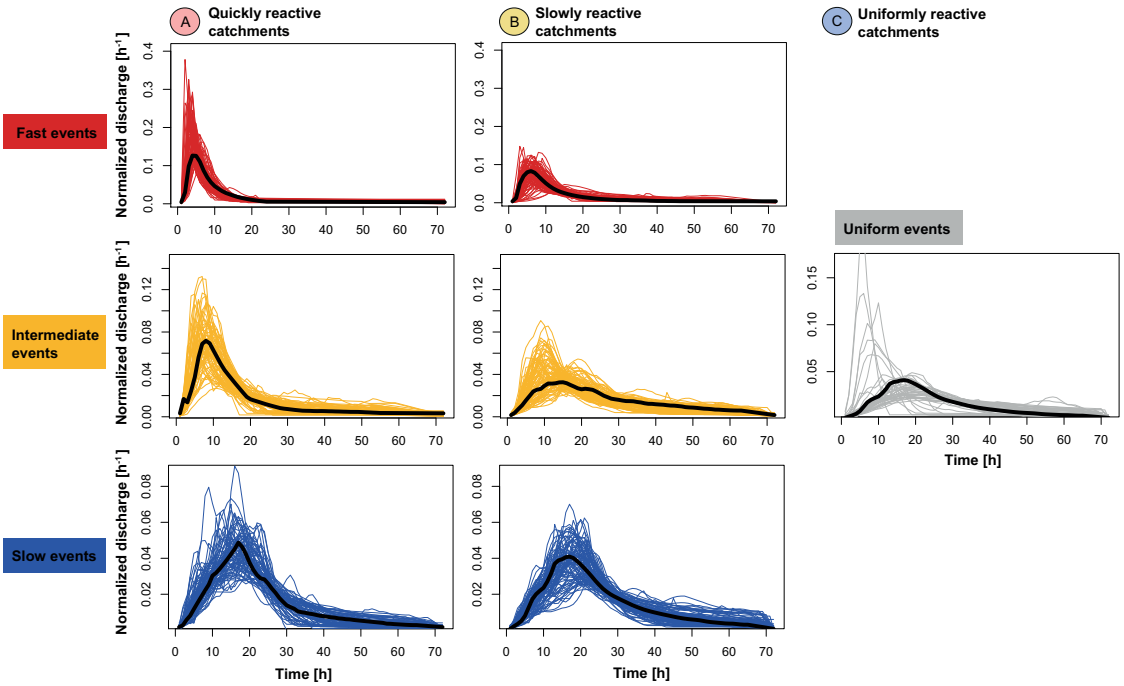



Figure 9: Representative hydrograph shapes of all catchments in a reactivity region (quickly reactive (A), slowly reactive (B), and uniformly reactive (C)). The median hydrograph shapes per region and event class (fast (red), intermediate (orange), and slow (blue)) are indicated in black. The scales of the $y$-axes are the same per event type but not across event types.

the Alps (see Figure 2). They are characterized by karstic geology with permeable rock, and a low network density. These catchment properties lead to a rather attenuated runoff reaction that leads to similar flood events independent of the triggering mechanism and antecedent wetness. Catchments with a generally fast runoff response, independent of the event type, built the quickly reactive region. Events occurring in this region were characterized by rather high peak discharges but rather low hydrograph volumes for all three event types. The catchments belonging to the quickly reactive region mainly lie in the Swiss Plateau. They were characterized by rather impermeable rocks (based on geological map of Switzerland (*Bundesamt für Statistik*, 2003)), and a high network density (based on the river network of Switzerland (*Swisstopo*, 2017)). These characteristics contributed to a fast runoff reaction. Finally, catchments with a generally slow or delayed runoff response formed the slowly reactive region. The events occurring in these catchments generally showed higher volumes but lower peak discharges compared to events occurring in the catchments of the quickly reactive region. Compared to the catchments in the quickly reactive region, they have more permeable rocks and lower network densities, which lead to an attenuation of the runoff events.

The distinct catchment characteristics of quickly, slowly, and uniformly reactive catchments allow for the establishment of a classification rule using tree-based algorithms such as random forest, bagging, or boosting (*James et al.*, 2013), which can be used to assign ungauged catchments to one of the reactivity classes. As a consequence, the reactivity regions could provide a useful basis for the regional estimation of hydrological model parameters, regional frequency analyses, the construction of design hydrographs in ungauged catchments, and other applications.

## 5.3 Perspectives

The clustering approach based on functional data proposed in this study can be adjusted to new regions or contexts. Here, we discuss four potential fields of future application: 1) application in other geographical regions and on other temporal scales, 2) Use of reactivity regions in design hydrograph construction, 3) application in a climate impact context, and 4) application in a fuzzy clustering setting.

1. **Application in other geographical regions and on other temporal scales:** The func-

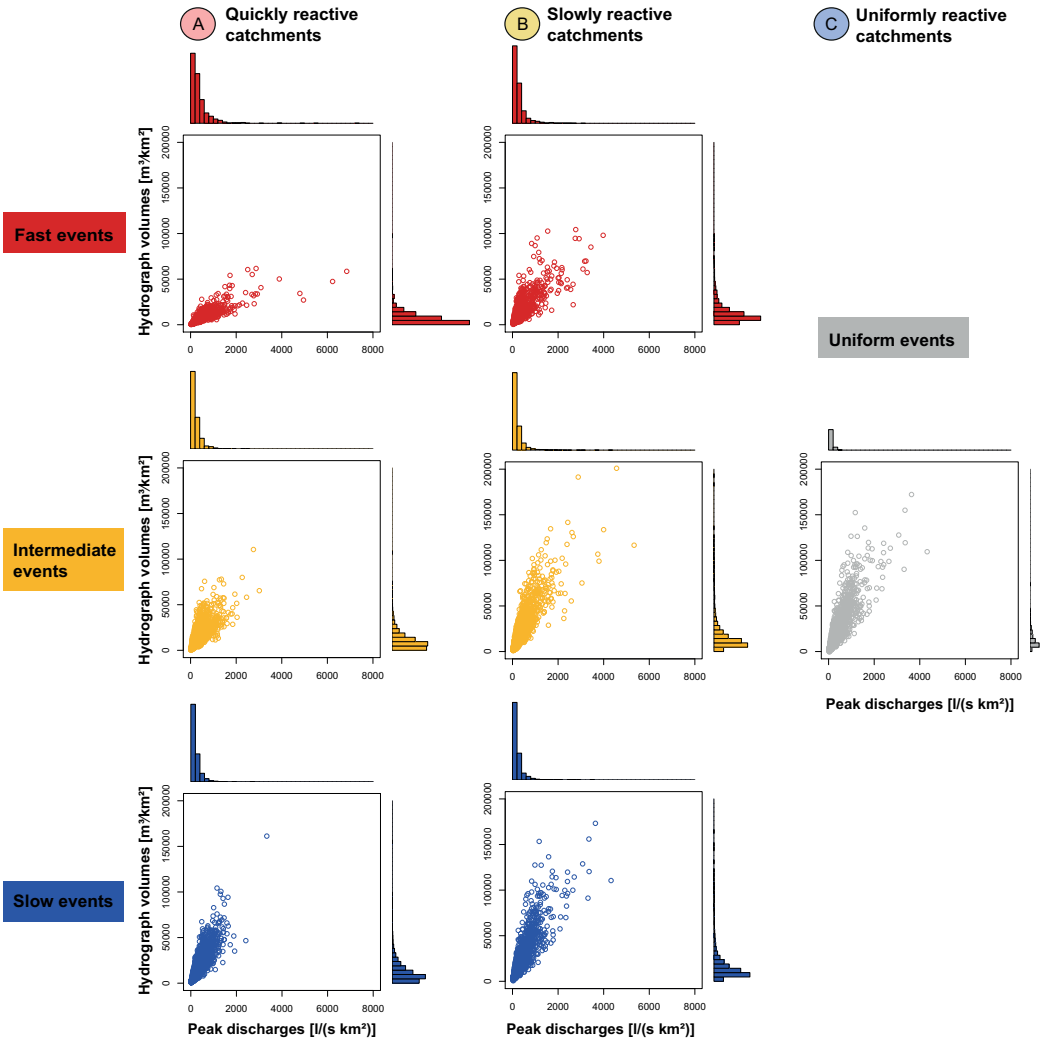



Figure 10: Specific peak discharges [l/(s km$^2$)] and hydrograph volumes [m$^3$/km$^2$] for the event types in the three reactivity regions (quickly reactive (A), slowly reactive (B), and uniformly reactive (C)). The two variables are jointly displayed as a scatter plot and separately in a histogram. The scales of the histograms and scatterplots are the same across all event types and reactivity regions.

tional approach to cluster flood hydrograph shapes for the identification of catchment-specific sets of hydrograph shapes and the subsequent clustering of these sets to identify reactivity regions has been developed and tested based on a set of 163 Swiss catchments with hourly flow series. The applicability of the approach is not limited to this geographical region but can be applied in other regions with similar data availability and catchments with similar physiographical and hydrological characteristics. However, its applicability might be limited to humid catchments since dry catchments were not included in the dataset. The applicability of the approach is neither limited to the use of hourly flow series but could be used on daily series if the catchments and their reaction times are larger than in Swiss catchments. The number of B-Splines used for the transformation of hydrograph shapes into FD might have to be increased in the case of more irregular flood hydrograph shapes than those observed in medium-size Swiss catchments. In addition, the number of shape clusters necessary to represent the hydrograph shape variability within a catchments might need to be adjusted. And finally, the number of reactivity regions might need to be adapted according to the variability of runoff and flood regimes in the region of interest.

2. **Use of reactivity regions in design hydrograph construction:** One potential field of

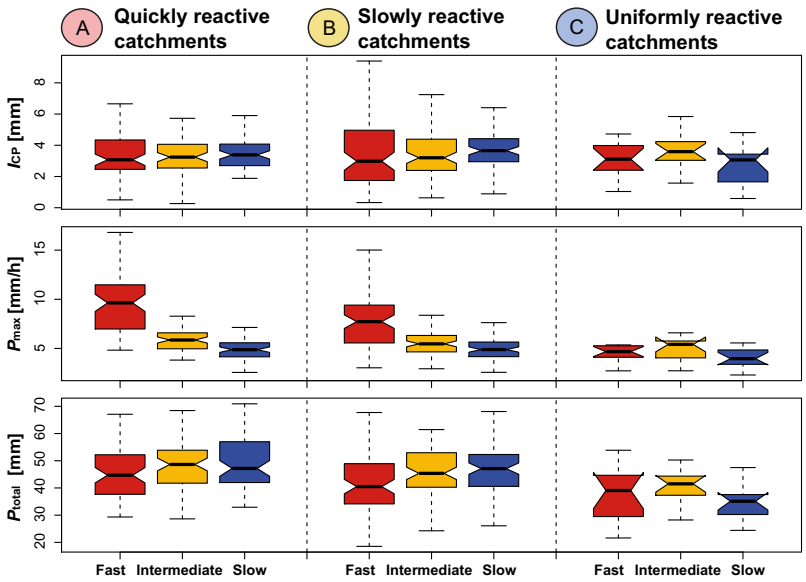



Figure 11: Event conditions for the three event types fast, intermediate, and slow in the three regions quickly reactive (A), slowly reactive (B), and uniformly reactive (C): Antecedent wetness ($I_{\mathrm{CP}}$), maximum hourly precipitation intensity ($P_{\mathrm{max}}$), and total event precipitation ($P_{\mathrm{total}}$). The notches of the boxplots indicate the confidence intervals around the medians.

application of the reactivity regions is design hydrograph construction in ungauged catchments. For this, the design hydrograph construction procedure proposed by *Brunner et al.* (2017) for gauged catchments can be adapted and combined with an index flood approach (*Dalrymple*, 1960; *Hosking and Wallis*, 1997). The design hydrograph construction approach is based on a bivariate flood frequency analysis, which considers the dependence between peak discharges and hydrograph volumes, and models the hydrograph shape via a probability density function (*Yue et al.*, 2002). Its combination with an index flood approach, where bivariate quantiles are estimated using pooled event data from several catchments, allows for an extension to ungauged catchments. The data pools can be formed by peak discharges and hydrograph volumes corresponding to the events belonging to the three event type classes within a reactivity region. The design variable estimates for peak discharges and hydrograph volumes of each of these classes can then be used together with the representative hydrograph shape of the corresponding event class to construct synthetic design hydrographs for a certain return period. An ungauged catchment can be attributed to one of the reactivity classes via a classification rule. The construction of synthetic design hydrographs in ungauged catchments can subsequently be based on the data of the reactivity class it is assigned to.

3. **Application in a climate impact context:** Under climate change, new flood types might develop and existing types might change (*Köplin et al.*, 2014; *Turkington et al.*, 2016) since different atmospheric circulation patterns favor different flood types (*Nied et al.*, 2017). The hydrograph clustering approach proposed here might be useful in a climate impact context for the detection of new flood types and changes in the proportion of occurrence of different flood types. This is important for future flood risk management since the flood type is closely linked to the spatial flood extent, temporal flood progression, and flood magnitudes (*Nied et al.*, 2014).

4. **Application in a fuzzy clustering setting:** *Rao and Srinivas* (2006) pointed out that most catchments only partly resemble other catchments in the region they have been assigned to. This is also the case when using the clustering approach proposed in this study and the assignment of a catchment to one group or another is not necessarily straightforward. Similarly, *Sikorska et al.* (2015) showed that flood events usually show partial memberships to more than one flood type. The clustering approach based on functional data proposed here could be transferred into a fuzzy setting (*Rao and Srinivas*, 2006; *Tokushige et al.*, 2007; *Sikorska et al.*, 2015) that would allow for catchments and hydrograph shapes with partial or distributed memberships to more

than one region or event type.

# 6 Conclusions

Flood hydrograph shapes contain a wealth of information that can be used for the identification of regions with a similar flood behavior but is insufficiently exploited by traditional multivariate clustering procedures. FD analysis makes much better use of the valuable information stored in a hydrograph and was found to be a useful tool for both clustering hydrograph shapes and the identification of flood reactivity regions. The clustering of hydrograph shapes within a catchment showed that three hydrograph shapes are sufficient to describe the hydrograph shape variability within a catchment. The sets of representative hydrograph shapes within a catchment were successfully used to establish regions with a similar flood behavior not only in terms of hydrograph shape but also hydrograph magnitude regarding peak discharges and hydrograph volumes. The clustering approach for hydrograph shapes and its use for the identification of flood reactivity regions has many potential fields of application. On the one hand, the reactivity regions could be used in regionalization studies, regional flood frequency analyses, and for design hydrograph construction in ungauged basins. On the other hand, the clustering approach could be applied to other geographical and climatological regions, transferred to a fuzzy clustering setting allowing for several partial group memberships, or be used in a climate impact context. The clustering approach for hydrograph shapes is a flexible and promising approach and has the potential to exploit the process information stored in flood hydrograph shapes.

# Acknowledgement

We thank the Federal Office for the Environment (FOEN) for funding the project (contract 13.0028.KP / M285-0623) and for providing runoff measurement data. We also thank MeteoSwiss for providing precipitation data. The data used in this study are available upon order from the FOEN and MeteoSwiss. For the hydrological data of the federal stations, the order form under `http://www.bafu.admin.ch/wasser/13462/13494/15076/index.html?lang=de` can be used. The hydrological data of the cantonal stations can be ordered from the respective cantons. The meteorological data can be ordered via `https://shop.meteoswiss.ch/index.html`. We thank the associate editor and the three reviewers for their constructive comments.

# References

Abraham, C., P. A. Cornillon, E. Matzner-Lober, and N. Molinari (2003), Unsupervised Curve Clustering using B-Splines, *Scandinavian Journal of Statistics*, *30*, 1–15.

Acreman, M. C., and C. D. Sinclair (1986), Classification of drainage basins acording to their physical characteristics; an application for flood frequency analysis in Scotland, *Journal of Hydrology*, *84*, 365–380, doi:10.1016/0022-1694(86)90134-4.

Ali, G., D. Tetzlaff, C. Soulsby, J. J. McDonnell, and R. Capell (2012), A comparison of similarity indices for catchment classification using a cross-regional dataset, *Advances in Water Resources*, *40*, 11–22, doi:10.1016/j.advwatres.2012.01.008.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (2013), *Runoff prediction in ungauged basins*, 465 pp., Cambridge University Press, Cambridge.

Boscarello, L., G. Ravazzani, A. Cislaghi, and M. Mancini (2016), Regionalization of flow-duration curves through catchment classification with streamflow signatures and physiographic-climate indices, *Journal of Hydrologic Engineering*, *21*(3), doi:10.1061/(ASCE)HE.1943-5584.0001307.

Brunner, M. I., D. Viviroli, A. E. Sikorska, O. Vannier, A.-C. Favre, and J. Seibert (2017), Flood type specific construction of synthetic design hydrographs, *Water Resources Research*, *53*, doi:10.1002/2016WR019535.

Bundesamt für Statistik (2003), Geodaten der Bundesstatistik.

Bundesamt für Umwelt BAFU, and Eidg. Forschungsanstalt WSL (2012), Faktenblatt LFI4 Viertes Schweizerisches Landesforstinventar LFI4 (2009-2011) - Zwischenergebnisse, *Tech. rep.*, BAFU und WSL, Birmensdorf ZH.

Burn, D. H., and D. B. Boorman (1992), Catchment classification applied to the estimation of hydrological parameters at ungauged catchments, *Tech. rep.*, Institute of Hydrology, Wallingford, Oxfordshire.

Castellarin, A., D. H. Burn, and A. Brath (2001), Assessing the effectiveness of hydrological similarity measures for flood frequency analysis, *Journal of Hydrology*, *241*(3), 270–285, doi:10.1016/S0022-1694(00)00383-8.

Chebana, F., S. Dabo-Niang, and T. B. M. J. Ouarda (2012), Exploratory functional flood frequency analysis and outlier detection, *Water Resources Research*, *48*(4), W04,514, doi:10.1029/2011WR011040.

Cuevas, A. (2014), A partial overview of the theory of statistics with functional data, *Journal of Statistical Planning and Inference*, *147*, 1–23, doi:10.1016/j.jspi.2013.04.002.

Cuevas, A., M. Febrero, and R. Fraiman (2007), Robust estimation and classification for functional data via projection-based depth notions, *Computational Statistics*, *22*(3), 481–496, doi:10.1007/s00180-007-0053-0.

Dalrymple, T. (1960), Flood-Frequency Analyses, *Geological survey water supply paper 1543-A*, p. 80.

Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall (2012), Merkblatt DWA-M 552, *Tech. rep.*, DWA, Hennef, Germany.

Diezig, R., and R. Weingartner (2007), Hochwasserprozesstypen — Schlüssel zur Hochwasserabschätzung, *Wasser und Abfall*, *4*, 18–26.

Eckhardt, K. (2005), How to construct recursive digital filters for baseflow separation, *Hydrological Processes*, *19*, 507–515, doi:10.1002/hyp.5675.

Fraiman, R., and G. Muniz (2001), Trimmed means for functional data, *Test*, *10*(2), 419–440.

Gordon, A. (1999), *Classification*, 2nd ed., 256 pp., Chapman & Hall/CRC, Boca Raton.

Hailegeorgis, T. T., and K. Alfredsen (2017), Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway, *Journal of Hydrology: Regional Studies Journal of Hydrology*, *9*, 104–126, doi:10.1016/j.ejrh.2016.11.004.

Hannah, D. M., B. P. G. Smith, A. M. Grunell, and G. R. McGregor (2000), An approach to hydrograph classification, *Hydrological Processes*, *14*, 317–338.

Höllig, K., and J. Hörner (2013), *Approximation and modeling with B-splines*, 214 pp., Society for industrial and applied mathematics, Philadelphia.

Hosking, J. R. M., and J. R. Wallis (1997), *Regional frequency analysis*, 238 pp., Cambridge University Press, Cambridge, doi:10.1017/CBO9780511529443.

Ilorme, F., and V. W. Griffis (2013), A novel procedure for delineation of hydrologically homogeneous regions and the classification of ungauged sites for design flood estimation, *Journal of Hydrology*, *492*, 151–162, doi:10.1016/j.jhydrol.2013.03.045.

Jacques, J., and C. Preda (2014), Model-based clustering for multivariate functional data, *Computational Statistics & Data Analysis*, *71*, 92–106, doi:10.1016/j.csda.2012.12.004.

Jamaludin, S. (2016), Streamflow Profile Classification using Functional Data Analysis : A Case Study on the Kelantan River Basin, in *The 3rd ISM international statistical conference*, vol. 1842, pp. 1–11, doi:10.1063/1.4982836.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2013), *An introduction to statistical learning. With applications in R.*, 418 pp., Springer, New York, doi:10.1007/978-1-4614-7138-7.

Köplin, N., B. Schädler, D. Viviroli, and R. Weingartner (2014), Seasonality and magnitude of floods in Switzerland under future climate change, *Hydrological Processes*, *28*(4), 2567–2578, doi:10.1002/hyp.9757.

Lang, M., T. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling, *Journal of Hydrology*, *225*, 103–117.

McDonnell, J. J., and R. Woods (2004), On the need for catchment classification, *Journal of Hydrology*, *299*, 2–3, doi:10.1016/j.jhydrol.2004.09.003.

Merleau, J., L. Perreault, J.-F. Angers, and A.-C. Favre (2007), Bayesian modeling of hydrographs, *Water Resources Research*, *43*, W10,432, doi:10.1029/2006WR005376.

Merz, R., and G. Blöschl (2003), A process typology of regional floods, *Water Resources Research*, *39*(12), 1340, doi:10.1029/2002WR001952.

Merz, R., and G. Blöschl (2009), Process controls on the statistical flood moments - a data based analysis, *Hydrological Processes*, *23*(5), 675–696, doi:10.1002/hyp.7168.

MeteoSwiss (2013), Documentation of MeteoSwiss Grid-Data Products Hourly Precipitation Estimation through Raingauge-Radar : CombiPrecip, *Tech. rep.*, MeteoSwiss, Zurich.

Nied, M., T. Pardowitz, K. Nissen, U. Ulbrich, Y. Hundecha, and B. Merz (2014), On the relationship between hydro-meteorological patterns and flood types, *Journal of Hydrology*, *519*, 3249–3262, doi:10.1016/j.jhydrol.2014.09.089.

Nied, M., K. Schröter, S. Lüdtke, V. D. Nguyen, and B. Merz (2017), What are the hydro-meteorological controls on flood characteristics?, *Journal of Hydrology*, *545*, 310–326, doi:10.1016/j.jhydrol.2016.12.003.

Oudin, L., A. Kay, V. Andréassian, and C. Perrin (2010), Are seemingly physically similar catchments truly hydrologically similar?, *Water Resources Research*, *46*, W11,558, doi:10.1029/2009WR008887.

Pilgrim, D. H. (1986), Bridging the Gap Between Flood Research and Design Practice, *Water Resources Research*, *22*(9), 165–176.

Prinzio, M. D., A. Castellarin, and E. Toth (2011), Data-driven catchment classification: application to the pub problem, *Hydrology and Earth System Sciences*, *15*, 1921–1935, doi:10.5194/hess-15-1921-2011.

Ramsay, J., and B. Silverman (2005), *Functional data analysis*, 426 pp., Springer, New York.

Ramsay, J., H. Wickham, S. Graves, and G. Hooker (2014), Package 'fda': Functional Data Analysis.

Ramsay, J. O., and B. W. Silverman (2002), *Applied functional data analysis: methods and case studies*, 190 pp., Springer, New York, doi:10.1007/b98886.

Rao, A. R., and V. V. Srinivas (2006), Regionalization of watersheds by fuzzy cluster analysis, *Journal of Hydrology*, *318*, 57–79.

Requena, A. I., F. Chebana, and T. B. M. J. Ouarda (2017), Heterogeneity measures in hydrological frequency analysis: review and new developments, *Hydrol. Earth Syst. Sci*, *21*, 1651–1668, doi:10.5194/hess-21-1651-2017.

Rousseeuw, P. J. (1987), Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, *Jounal of Computational and Applied Mathematics*, *20*, 53–65.

Sikorska, A. E., D. Viviroli, and J. Seibert (2015), Flood type classification in mountainous catchments using crisp and fuzzy decision trees, *Water Resources Research*, *51*(10), 7959–7976, doi:10.1002/2015WR017326.

Smakhtin, V. Y., and B. Masse (2000), Continuous daily hydrograph simulation using duration curves of a precipitation index, *Hydrological Processes*, *14*, 1083–1100.

Swisstopo (2017), Swiss Map Vector 25 BETA.

Ternynck, C., M. Ali, B. Alaya, F. Chebana, S. Dabo-Niang, and T. B. M. J. Ouarda (2016), Streamflow hydrograph classification using functional data analysis, *American Meteorological Society*, *17*, 327–344, doi:10.1175/JHM-D-14-0200.1.

Tokushige, S., H. Yadohisa, and K. Inada (2007), Crisp and fuzzy k-means clustering algorithms for multivariate functional data, *Computational Statistics*, *22*(1), 1–16, doi:10.1007/s00180-006-0013-0.

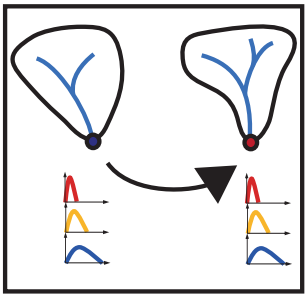
Turkington, T., K. Breinl, J. Ettema, D. Alkema, and V. Jetten (2016), A new flood type classification method for use in climate change impact studies, *Weather and Climate Extremes*, *14*, 1–16, doi:10.1016/j.wace.2016.10.001.

Ward, J. H. (1963), Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, *58*(301), 236–244, doi:10.1080/01621459.1963.10500845.

Webster, R., and M. A. Oliver (2007), *Geostatistics for environmental scientists*, Statistics in practice, 330 pp., John Wiley & Sons Ltd., Chichester.

Yue, S., T. Ouarda, B. Bobée, P. Legendre, and P. Bruneau (2002), Approach for describing statistical properties of flood hydrograph, *Journal of Hydrologic Engineering*, *7*(2), 147–153, doi:10.1061/(ASCE)1084-0699(2002)7:2(147).

# Paper VI



Advances in Water Resources 112 (2018) 235–244

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

# Representative sets of design hydrographs for ungauged catchments: A regional approach using probabilistic region memberships

Manuela Irene Brunner[*,a,b], Jan Seibert[a,c], Anne-Catherine Favre[b]

[a] Department of Geography, University of Zurich, Zurich, Switzerland
[b] Université Grenoble-Alpes, CNRS, IRD, IGE, Grenoble INP, Grenoble, France
[c] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

ARTICLE INFO

*Keywords:*
Classification
Random forest
Homogeneous regions
Regionalization
Mixtures
Floods

ABSTRACT

Traditional design flood estimation approaches have focused on peak discharges and have often neglected other hydrograph characteristics such as hydrograph volume and shape. Synthetic design hydrograph estimation procedures overcome this deficiency by jointly considering peak discharge, hydrograph volume, and shape. Such procedures have recently been extended to allow for the consideration of process variability within a catchment by a flood-type specific construction of design hydrographs. However, they depend on observed runoff time series and are not directly applicable in ungauged catchments where such series are not available. To obtain reliable flood estimates, there is a need for an approach that allows for the consideration of process variability in the construction of synthetic design hydrographs in ungauged catchments. In this study, we therefore propose an approach that combines a bivariate index flood approach with event-type specific synthetic design hydrograph construction. First, regions of similar flood reactivity are delineated and a classification rule that enables the assignment of ungauged catchments to one of these reactivity regions is established. Second, event-type specific synthetic design hydrographs are constructed using the pooled data divided by event type from the corresponding reactivity region in a bivariate index flood procedure. The approach was tested and validated on a dataset of 163 Swiss catchments. The results indicated that 1) random forest is a suitable classification model for the assignment of an ungauged catchment to one of the reactivity regions, 2) the combination of a bivariate index flood approach and event-type specific synthetic design hydrograph construction enables the consideration of event types in ungauged catchments, and 3) the use of probabilistic class memberships in regional synthetic design hydrograph construction helps to alleviate the problem of misclassification. Event-type specific synthetic design hydrograph sets enable the inclusion of process variability into design flood estimation and can be used as a compromise between single best estimate synthetic design hydrographs and continuous simulation studies.

## 1. Introduction

Classical design flood estimation has been focusing on the univariate analysis of peak discharges even though other event characteristics such as hydrograph volume and shape are equally important for hydraulic design tasks involving storage (Pilgrim, 1986). More recent flood estimation procedures allow the representation of both the magnitude and the shape of an event through synthetic design hydrographs (SDHs). Synthetic design hydrographs provide a more complete picture of the flood behavior of a catchment than classical approaches. SDHs include event-based approaches using event rainfall as input (Grimaldi et al., 2012; Rogger et al., 2012) and statistical approaches using observed runoff in bivariate flood frequency analyses (Brunner et al., 2017b). However, as the classical approaches, they neglect the variability of flood events within a catchment caused by different processes, which are mirrored by various flood types, such as flash floods, short-rain-, long-rain-, and rain-on-snow floods (Merz and Blöschl, 2003). To overcome this deficiency, (Brunner et al., 2017b) proposed an approach for the construction of a set of flood-type specific SDHs. The approach splits the flood sample into four subsets, one for each flood type, and uses each of these samples to construct a flood-type specific design hydrograph. The shape of the design hydrographs is modeled by a probability density function (PDF) while the magnitude of the event is modeled via a bivariate frequency analysis taking into account the dependence between peak discharges and hydrograph volumes via a copula model. This ensemble-based SDH construction

* Corresponding author.
*E-mail address:* manuela.brunner@geo.uzh.ch (M.I. Brunner).

https://doi.org/10.1016/j.advwatres.2017.12.018
Received 26 October 2017; Received in revised form 22 December 2017; Accepted 22 December 2017
Available online 24 December 2017

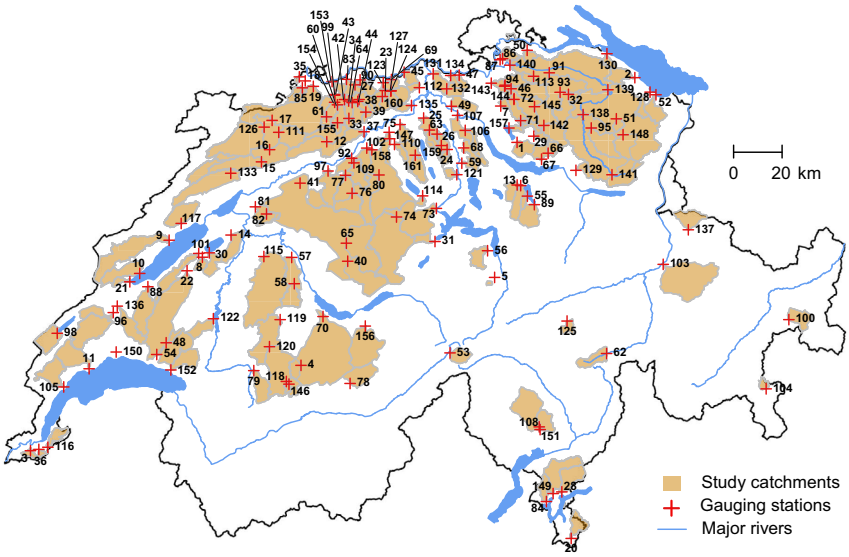



**Fig. 1.** 163 study catchments in Switzerland.

approach takes into account process variability but is based on observed flood events and cannot be easily transferred to ungauged catchments where no runoff information is available. There, design floods have traditionally been estimated by a regional index flood approach focusing on peak discharges.

The index flood approach consists of two main steps. In a first step, regions with a similar flood behavior are delineated. In a second step, the data within these similar regions are used for regional flood frequency analysis. Hydrologically similar regions are often delineated based on hydrological catchment characteristics or runoff signatures (Burn and Boorman, 1992) since physiographical and climatological catchment similarity do often not correspond to hydrological similarity (Ali et al., 2012; Oudin et al., 2010). Brunner et al. (2017a) suggested to use entire hydrograph shapes in the delineation of homogeneous regions with the argument that those provide more information on the flood behavior of a catchment than statistical measures of individual hydrograph characteristics. In this previous study, it was shown that flood reactivity regions can be delineated by characterizing each catchment in the dataset by a set of three representative hydrograph shapes: a fast, an intermediate, and a slow shape. Grouping catchments with similar sets of representative hydrograph shapes delineates regions which are similar in terms of their flood behavior. Such regions were shown to have a hydro-meteorological meaning and are potentially useful in regional flood frequency analysis.

Regional frequency analysis is often done using the index flood approach which was proposed by Dalrymple (1960) for annual maxima series and later extended to partial duration (peak-over-threshold) series (Madsen et al., 1997). It assumes that frequency distributions at different sites within a region are identical apart from a scale factor. It describes a local quantile estimate $Q_i(F)$ as the product of an index flood ($\mu_i$) and a regional growth curve ($q(F)$) estimated based on the data at $N$ sites as:

$$Q_i(F) = \mu_i q(F) \quad i = 1, \ldots, N. \tag{1}$$

The index flood can be any location measure of the at-site distribution but is often taken to be its mean. The regional growth curve is a dimensionless quantile function computed based on dimensionless regional data, which are obtained by dividing the observed flood event data by the index flood. Regional analysis using the index flood approach yields more accurate quantile estimates than at-site analysis even if a region is heterogeneous. It was thus found to be a robust and efficient estimation procedure (Lettenmaier et al., 1987; Madsen et al., 1997).

The classical index flood procedure focuses on peak discharges. Requena et al. (2016) therefore proposed an approach for a multivariate regional index approach that allows for the consideration of more than one design hydrograph characteristic, e.g. peak discharge and hydrograph volume. While such a bivariate regional approach allows the joint consideration of peak discharges and hydrograph volumes, neither hydrograph shape, nor process variability can be considered. To our knowledge, no methodology has so far been proposed for the regional construction of event-type specific sets of SDHs in ungauged catchments. The aim of this study was therefore to propose an approach that allows for the construction of SDHs in ungauged catchments which on the one hand jointly represents the magnitude and shape of an event and on the other hand allows for the consideration of process variability. We here propose an approach that delineates regions of a similar flood reactivity using the approach by Brunner et al. (2017a), applies the bivariate index flood approach proposed by Requena et al. (2016) within these flood reactivity regions, and uses the resulting design variable pairs in the SDH construction approach proposed by Brunner et al. (2017b). Instead of flood types, we use the three event types fast, intermediate, and slow in design hydrograph construction Brunner et al. (2017a). The three event-type specific SDHs together form a set of design hydrographs which considers the process variability within an ungauged catchment.

Our research more specifically addresses the following research questions:

1. How can an ungauged catchment best be assigned to one of the flood reactivity regions?
2. How can event-type specific SDH sets for ungauged catchments be constructed?
3. Can probabilistic class memberships be used in regional SDH construction to reduce the problem of misclassification?

## 2. Data

This analysis used runoff and catchment characteristics data from 163 Swiss catchments (Fig. 1) with a wide range of catchment characteristics and flood behaviors. The selected catchments have hourly flow series of at least 20 years in duration ranging up to 53 years. In these catchments, runoff is neither altered by regulated lakes upstream or inland canals nor by urbanized areas. The catchments are small to medium-size (6 to 1800 km$^2$), situated at mean elevations between 400 and 2600 m.a.s.l., and have no or only a minor glacier coverage.

The basis for the analysis was samples of flood events extracted from the runoff time series of the 163 study catchments. To sample flood events, we used a peak-over-threshold approach based on the procedure proposed by Lang et al. (1999). The threshold for the peak discharge was chosen iteratively to fulfill a target condition of four events per year on average which is a trade-off between maximizing the information content in the sample and keeping the assumption of independence between events. For each of these events, sampled according to the flood peaks, the flood volume was determined over the actual event duration. The actual duration was determined as the time between the onset of the event (defined as the time where discharge first exceeds 0.05 times the peak discharge) and the end of the event (discharge falls below 0.05 times the peak discharge). The baseflow was separated from the direct flow using a recursive digital filter (Eckhardt, 2005). The resulting direct flow component of the hydrographs was then normalized so that the volume of the modified hydrographs was equal to one. This was done by dividing the ordinate of each hydrograph by the volume $V$. In the remainder of this paper, we refer to these normalized hydrographs as *hydrograph shapes*. To make the sampled event hydrographs comparable, they were brought to a length of 72 hours by appending hours with the minimum discharge until the series consisted of 72 values. We refer the reader to Brunner et al. (2017b) for a more detailed description of the flood sampling and baseflow separation procedures.

## 3. Methods

In the regional approach for the construction of event-type specific sets of SDHs, we assigned an ungauged catchment to a flood reactivity region and subsequently used the data from this reactivity region for the estimation of a set of SDHs representing three event types comprising a fast, an intermediate, and a slow event. The regional approach consists of three main steps: (1) Delineation of reactivity regions by clustering catchments using their catchment-specific sets of representative hydrograph shapes, (2) Establishment of a classification rule allowing the attribution of an ungauged catchment to one of the reactivity regions, (3) Construction of a set of event-type specific design hydrographs using I) region specific hydrograph shape sets and II) event magnitudes estimated by event type in a bivariate index flood approach. The individual steps of the methodology are illustrated in Fig. 2 and described in more detail in the following paragraphs.

### 3.1. Step 1: Delineation of regions with similar flood reactivity

Brunner et al. (2017a) showed that the hydrograph shape variability within Swiss catchments can be summarized by a set of three representative hydrograph shapes: a fast, an intermediate, and a slow hydrograph. In this previous study, we represented normalized hydrograph shapes as functional data, i.e., as continuous functions, by projecting them on a set of four basis splines (B-spline) bases. A spline function is determined by the order of the polynomial segments and the number and placement of knots. The number of knots determines the ability of spline functions to represent sharp features in a curve and the knots can be placed such that they are denser in areas with stronger variations than in smooth areas (Höllig and Hörner, 2013). The sets of four coefficients, one per B-spline base, were used in a $k$-means clustering algorithm (Gordon, 1999) to identify clusters of similar hydrograph shapes. Three shape clusters were found to well explain the shape variability within Swiss catchments. Each cluster was summarized by its median hydrograph shape. The three median shapes together formed the set of representative hydrograph shapes consisting of a fast, an intermediate, and a slow hydrograph shape. The fast event type is characterized by both steep rising and recession limbs. The intermediate event type is characterized by a rather steep rising but a slow recession limb, and the slow event type is characterized by both slow rising and recession limbs (Brunner et al., 2017a). Brunner et al. (2017a) then used these catchment-specific sets of representative hydrograph shapes for the identification of flood reactivity regions i.e., regions that were similar in terms of their representative hydrograph shape sets. We modified the approach proposed by Brunner et al. (2017a) in the following two ways: 1) The clustering was done using the representative hydrograph sets of all catchments without separating a uniformly reactive catchment as done in Brunner et al. (2017a) because this was found to be a disadvantage for the identification of a classification rule in Step 2. 2) We delineated four instead of three regions with a similar flood reactivity using the functional representation of catchment-specific sets of hydrograph shapes as an input for the hierarchical clustering algorithm. Four regions were found to be better since the hierarchical clustering tree showed a clear symmetry and cutting it at three or five clusters would not have been sensible. We called these regions A to D. Region A is characterized by a quick runoff reaction, i.e., all three representative hydrograph shapes showed a rather quick response compared to the representative hydrograph shapes in catchments belonging to the other three regions. Regions B and C were characterized by an intermediate or slow reactivity respectively. Region D showed a rather uniform reaction to rainfall input (i.e., the fast, intermediate, and slow shapes were difficult to distinguish). The runoff behavior of catchments within a region was summarized by a set of median event-type specific hydrograph shapes (median shape sets). The cluster memberships of the 163 study catchments were used in Step 2 to identify a classification rule allowing for the assignment of an ungauged catchment to one of the four reactivity regions.

### 3.2. Step 2: Attribution of ungauged catchment to reactivity region

We established a relationship between the catchment characteristics and the region memberships of catchments to be able to attribute an ungauged catchment to one of the four reactivity regions based on catchment characteristics only. An initial comparison of several classification methods showed that random forest (Harrell, 2015; James et al., 2013) was the most suitable classification model in cross-validation. The set of explanatory variables used to fit the models consisted of the following seven weakly correlated catchment characteristics: catchment area, network density, $Y$-coordinate, soil topographic index, percentage area of karstic rocks, sunshine duration, and vapor pressure. Catchment area and network density were derived from the digital elevation model, the soil topographic index from the digital map of land surface characteristics (Eidgenössische Forschungsanstalt für Wald Schnee und Landschaft (WSL), 1999), the percentage area of karstic rock from a map focusing on groudwater resources (Bitterli et al., 2007), and sunshine duration and vapor pressure from gridded meteorological data provided by MeteoSwiss (MeteoSwiss, 2013). We used the random forest classification model to attribute ungauged catchments to one of the reactivity regions according to its best class membership and to compute probabilities of class memberships, which we hereafter refer to as *probabilistic* class memberships, for each of the four regions. Considering probabilistic class memberships allowed the alleviation of the problem of misclassification. Regions B and C were found to be difficult to distinguish when assigning a best class membership to a catchment. Assigning probabilistic class memberships and using them in SDH construction helped to take into consideration this source of uncertainty.

### 3.3. Step 3: Construction of a set of event-type specific SDHs

An ungauged catchment was assigned to one of the reactivity regions delineated in Step 1 via the random forest classification model established in Step 2. The construction of an event-type specific set of SDHs for this ungauged catchment was based on the pool of data of the reactivity region it was assigned to (best class membership). The approach was at a later stage extended to probabilistic region memberships as described under Step 2. We first focus on the construction of

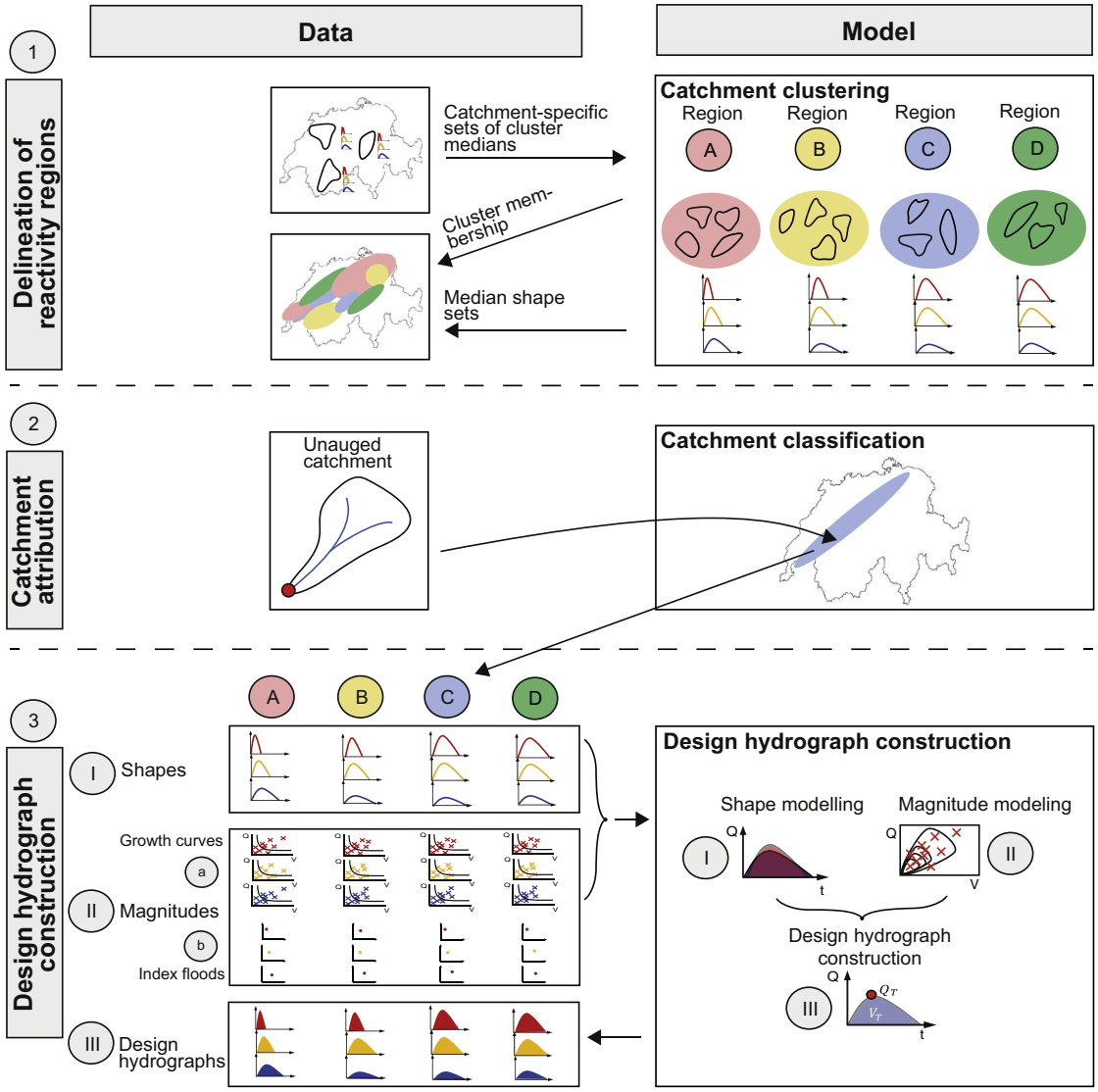



**Fig. 2.** Regional event-type specific SDH construction framework: (1) Delineation of reactivity regions, (2) Catchment attribution, (3) Design hydrograph construction. Design hydrographs (III) characterize a flood both in terms of hydrograph shape (I) and hydrograph magnitude (II). The magnitude was estimated using event-type specific index floods (b) and regional growth curves (a).

event-type specific sets of SDHs using the best class membership. The event-type specific construction procedure was based on the SDH construction procedure proposed by Brunner et al. (2017b) for different flood types (Merz and Blöschl, 2003; Sikorska et al., 2015). Instead of differentiating flood types, we here distinguished between the three event types fast, intermediate, and slow, which cover a large part of shape variability within a catchment and were used in Section 3.1 for the delineation of reactivity regions. A set of three SDHs was constructed by combining shape estimates represented by a probability density function with magnitude estimates represented by the bivariate design variable quantiles peak discharge and hydrograph volume. The basic idea of the procedure is to do regional flood frequency analysis within the reactivity region of interest using a bivariate index flood approach (Requena et al., 2016) for each of the three event types (fast, intermediate, and slow) separately. This regional SDH construction approach consists of three steps (see Fig. 2): (I) Computation of design event shape, (II) Computation of design event magnitude by estimating a) a regional growth curve, and b) a local index flood, and (III) Construction of design hydrograph by combining magnitude and shape. These steps are described in detail in the following paragraphs.

#### 3.3.1. (I) Computation of design event shape

The design hydrograph shapes were estimated using the three representative hydrograph shapes (fast, intermediate, and slow) of the region the ungauged catchment was assigned to. The three hydrograph shapes were each fitted by a lognormal probability density function (PDF) (Yue et al., 2002), which was expressed in terms of their time to peak, peak discharge, and time base (Nadarajah, 2007; Rai et al., 2009). The lognormal PDF was chosen out of a selection of eight PDFs (normal, lognormal, Fréchet, Weibull, beta, gamma, inverse gamma, and logistic) whose fit to the representative hydrograph shapes was assessed via the Kling–Gupta and Nash–Sutcliffe efficiencies (Gupta et al., 2009).

#### 3.3.2. (II) Computation of design event magnitude

We formed pools of observed flood events for each event type within the predicted reactivity region. The region was characterized by three pools of data: fast events, intermediate events, and slow events. The individual data pools consisted of all the event-type specific events of the catchments within the corresponding reactivity region. The data pools consisted of dimensionless peak discharges (peak discharges normalized by mean catchment peak discharge [i.e., index flood peak]) and dimensionless hydrograph volumes (hydrograph volumes

normalized by mean catchment hydrograph volume [i.e., index hydrograph volume]) to make values from different catchments comparable. These pooled data were used in a bivariate flood frequency analysis to derive regional growth curves for peak discharges and hydrograph volumes (Requena et al., 2016). The bivariate frequency analysis was based on the marginal distributions of peak discharges ($Q_p$) and hydrograph volumes ($V$) and their dependence was modeled via a copula function.

a) **Regional growth curves:** The four regions were characterized by differences in $Q_p$ and $V$. $Q_p$ generally decreased from Region A to D while $V$ increased. $Q_p$ and $V$ not only differed between regions but also between the three event types within a catchment. Fast and intermediate events were generally characterized by higher magnitudes in terms of $Q_p$ and smaller magnitudes in terms of $V$ than slow events. We used the flexible five-parameter Wakeby distribution (Houghton, 1978; Griffiths, 1989) to model the marginal distributions of $Q_p$ and $V$ because this distribution can mimic the shapes of many commonly used skew distributions such as the extreme value distribution, the lognormal, or Pearson type III distributions. The Wakeby distribution provided a good fit to the data, which was confirmed by the Kolmogorov–Smirnov goodness-of-fit test (level of significance $\alpha = 0.05$). Fitting a distribution with five parameters was not a problem in our case since the sample size was sufficiently large when working with the pooled data. The Wakeby distribution ($F(x)$) is not explicitly defined. However, its inverse ($x(F)$) can be expressed by the five parameters $\alpha$, $\beta$, $\gamma$, $\lambda$, and $\xi$ (Hosking, 1986) as follows:

$$x(F) = \xi + \frac{\alpha}{\beta}[1 - (1 - F)^{\beta}] - \frac{\gamma}{\delta}[1 - (1 - F)^{-\delta}]. \quad (2)$$

When $\delta > 0$, the Wakeby distribution has a heavy upper tail and can give rise to data sets containing occasional high outliers. The upper-tail behavior of the Wakeby distribution is determined by the parameters $\gamma$ and $\delta$ unless $\gamma = 0$. As $F \rightarrow 1$, the density function of the Wakeby distribution is asymptotically equivalent to that of a generalized Pareto distribution (Hosking, 1986). Other distributions commonly used in flood frequency analysis such as the generalized extreme value distribution and the generalized Pareto distribution (Coles, 2001) did not provide a good fit to the data. Compared to the Wakeby distribution, their flexibility is limited by the lower number of parameters (three). The parameters of the Wakeby distribution were estimated for both the $Q_p$ and $V$ in each of the data pools using $L$-moments (Hosking and Wallis, 1997) and its parameters were found to differ between the data pools.The form of the dependence between $Q_p$ and $V$ did not strongly differ for the three event types within a region nor was it different between the four reactivity regions. It could therefore be modeled by the same copula family for all event types and regions. The dependence between $Q_p$ and $V$ was modeled using the elliptical Student-$t$ copula (Frahm et al., 2003), which is able to model lower and upper non null tail dependence as present in the data (assessed via the tail dependence estimator proposed by Schmid and Schmidt (2007)). We tested the suitability of several copula families: the independence copula, several Archimedean (Gumbel, Clayton, Joe, Frank, AMH, Hüsler-Reiss, Galambos, Farlie-Gumbel-Morgenstern (FGM), Tawn, Plackett, and survival Clayton) and two elliptical copulas (Normal, Student-$t$) to model the dependence between $Q_p$ and $V$ by computing the Cramér-von-Mises goodness-of-fit statistic (Genest and Favre, 2007). The elliptical copulas provided a better fit to the data than the Archimedean copulas when comparing their Cramér-von-Mises test statistics. The Student-$t$ copula was chosen instead of the Normal copula because it is able to model positive tail dependence (Frahm et al., 2003). In addition to the form of the dependence, the intensity of the dependence, measured by Kendall's tau and Spearman's rho, only slightly differed for the three event types within a region and between the four reactivity regions.The pair of dimensionless growth curves for $Q_p$ and $V$ was estimated using the marginal distributions of the variables represented by Wakeby distributions and the Student-$t$ copula for a joint return period of $T$ considering that both design variables are equally important (Brunner et al., 2016). We here focused on a return period of $T = 100$ years since it is often used in practice (Camezind-Wildi, 2005). However, the approach is not limited to this return period.

b) **Index floods:** The index floods were computed for both peak discharges and hydrograph volumes and for each of the regions and event types separately. The index flood peak and volume were predicted based on a Gamma generalized linear model (GLM) with a log-link, which does not give rise to a negative estimated response (Myers et al., 2010), using the same catchment characteristics as explanatory variables as used for establishing the classification rule (Section 3.2). The GLMs were fitted for each of the twelve ($4 \times 3$) data pools separately. We used a stepwise regression procedure (Harrell, 2015) to identify the model with the fewest explanatory variables still providing us with a useful model for index flood prediction. We found that the index flood peaks and volumes could be predicted by only a few explanatory variables, which were similar for both design variables ($Q_p$ and $V$). The most important explanatory variables across all regions and event types were catchment area and network density.

The regional growth curves were used together with the catchment specific index floods to obtain local design variable quantile estimates using Eq (1).

#### 3.3.3. (III) Construction of design hydrograph by combining magnitude and shape

The three estimated design variable pairs (Section 3.3.2) and the regional PDFs (Section 3.3.1) were used to construct three representative SDHs for an ungauged catchment. An SDH can be expressed as:

$$Q_T(t) = f(t)(V_T/D_T) + B, \quad (3)$$

where $f(t)$ represents the PDF used to model the shape of the hydrograph, $V_T$ and $D_T$ the design quantiles for hydrograph volume and duration for the return period $T$, and $B$ the baseflow component. $D_T$ can be derived as $D_T = f(t_p)(V_T/Q_T)$, where $t_p$ is the time to peak and $Q_T$ the design quantile for peak discharge. Baseflow was added via a mean event baseflow index, which was computed per reactivity region (independently of the event type), proportionally to the direct runoff (see Brunner et al. (2017b)).

The construction of event-type specific SDH sets using probabilistic class memberships proceeded similarly to the construction procedure described above when using the best class membership. It again differentiated between pools of fast, intermediate, and slow events. To compute the design event magnitude, we computed the regional growth curves for each of the four regions and used them to compute an averaged regional growth curve weighting the individual growth curves by the probabilities of membership to each of the corresponding regions. In addition, we predicted four index floods using the four GLMs for each of the regions and computed again a weighted average of these predictions using the probabilities of region membership as weights. The design event magnitude was finally computed by upscaling the weighted average of the regional growth curves with the weighted average of the index floods. The design event shapes were also computed as a weighted average from the shapes of the four regions using again the probabilities of region memberships as weights for the individual PDFs.

The event-type specific regional SDH procedure was applied both using the best class membership and the probabilistic class memberships. This resulted in two sets of regionally estimated SDHs.

### 3.4. Validation of approach

We compared these two sets of regionally estimated SDHs to SDHs computed using local runoff observations. The local estimation was done based on the event-type specific data within a catchment using the procedure proposed in Brunner et al. (2017b) for the construction of flood-type specific SDHs. This meant that three event-type specific SDHs were computed for each catchment. Local event-type specific SDHs were only constructed for event types with more than five observations. This prevented from completely unreliable estimates, however, every estimate obtained by a sample of less than twenty observations had to be considered to be not really reliable (Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall, 2012). This meant that even though the local estimates have been computed using observed data, they might not represent the true SDHs for the catchment under consideration. Nonetheless, we used them as a basis for validation. We computed the relative and the absolute relative error of the regional estimates compared to the local estimates for four hydrograph characteristics: peak discharge ($Q_p$), hydrograph volume ($V$), time to peak ($t_p$), and half-recession time ($t_{p05}$), i.e., the time from peak to where the recession curve falls back to half the peak discharge.

## 4. Results

### 4.1. Region assignment

The catchment-specific sets of representative hydrograph shapes were used to delineate regions with similar flood behaviors. Their clustering resulted in four clusters/regions with distinct flood-event reaction-times: Region A with catchments with a generally fast runoff reaction (see Fig. 3A, 30 catchments), Region B with catchments with a generally intermediate runoff reaction (see Fig. 3B, 45 catchments), Region C with catchments with a generally slow runoff reaction (see Fig. 3C, 58 catchments), and Region D with catchments with a generally rather uniform runoff reaction (see Fig. 3D, 30 catchments). The catchments belonging to Region A were mainly small catchments located in the Swiss Plateau, catchments belonging to Region D were mostly located in the Jura mountains, and regions B and C consisted of catchments in the Swiss Plateau and in Alpine regions.

Catchments in Region A were characterized by hydrograph sets with a steeper recession limb than the catchments in the regions B, C, and D (Fig. 4). The differences in hydrograph shapes between the four regions were largest for the fast event shapes, well visible for the intermediate event shapes, and rather weak for the slow event shapes. The reactivity regions not only differed in terms of their representative hydrograph shapes but also in terms of hydrograph magnitudes (Table 1). Events occurring in Region A were characterized by rather high peak discharges compared to flood volumes while events occurring in region D showed rather high volumes compared to peak discharges. Regions B and C lay somewhere in between.

An ungauged catchment can be assigned to one of the four reactivity regions via a classification rule. We found that the most suitable model to establish a classification rule was random forest, which had a misclassification error of 45% (see Fig. 5). This implies that an ungauged catchment will be attributed to the correct region with a probability of 55%. The classification error could mainly be explained by catchments attributed to Region B instead of C and vice versa (17%). These two clusters seemed to be difficult to distinguish using catchment characteristics, i.e., the probabilities of belonging to one or the other regions were quite similar. Some catchments were also attributed to Region C instead of D. Only a few catchments were attributed to a non-neighboring region (e.g. to Region C instead of A) (12%). The most important catchment characteristics for predicting the region membership of a catchment among the seven characteristics used were found to be: catchment area, network density, and location in space. Region A was characterized by small catchments with a rather high network density and low sunshine duration. These catchments were mainly located in the Swiss Plateau. On the contrary, region D was characterized by large catchments, with a low network density, high sunshine duration, and a high percentage area of karstic rock. These catchments were mainly located in the Jura Mountains and Northeastern Switzerland. The catchments belonging to regions B and C were medium-size, characterized by medium network densities, and sunshine durations. They were located in both the Swiss Plateau and the Alps and were not easy to attribute to one or the other region.

### 4.2. Regional event-type specific SDH sets

The previously established regions were used in regional flood frequency analysis to derive two sets of representative design hydrographs for ungauged catchments for a specified return period. We focused on the two return periods 10 and 100 years for illustration purposes since these return periods are often used in practice (Camezind-Wildi, 2005). As described in Section 3.3, we computed two regional sets of three hydrographs (fast, intermediate, and slow). The first set was computed using regional information of the region with the highest probability of membership (called best class membership). On the contrary, the second set was computed using regional information of the four regions weighted according to the probabilistic region memberships. Fig. 6 shows the two regionally estimated SDH sets (thin and thick bold lines) together with a locally estimated SDH set (dashed lines) and a catchment-specific SDH not distinguishing between event types computed using local observations (black dashed line) for a return period of 100 years. The two sets of regionally estimated SDHs compared well with the locally estimated SDHs in most catchments. Furthermore, the regionally estimated SDHs lay in the order of magnitude of the highest observed event of a catchment (grey lines).

The relative error of the four hydrograph characteristics peak discharge, hydrograph volume, time to peak, and half-recession time was roughly 50% for all the event types for both return periods considered (10 and 100 years) (Fig. 7). However, the variability of the relative errors across catchments was slightly lower for $T = 10$ than for $T = 100$. The relative errors were similar when using the best and the probabilistic class memberships. While the median relative error was not highly affected by using probabilistic memberships instead of the best class membership, the variability of the relative errors was clearly higher when applying the best class membership. The relative error of peak discharges was independent of the event type while it depended on the event type for the other three hydrograph characteristics. It was

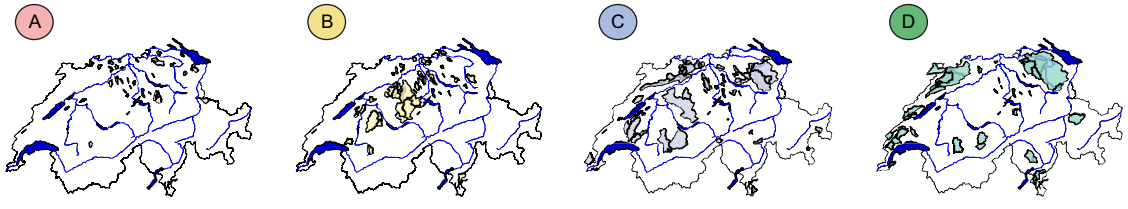



**Fig. 3.** Four reactivity regions A to D. The runoff reaction decreased from catchments in Region A with a generally fast reaction over catchments in Regions B and C to catchments in region D with a generally rather uniform runoff reaction.

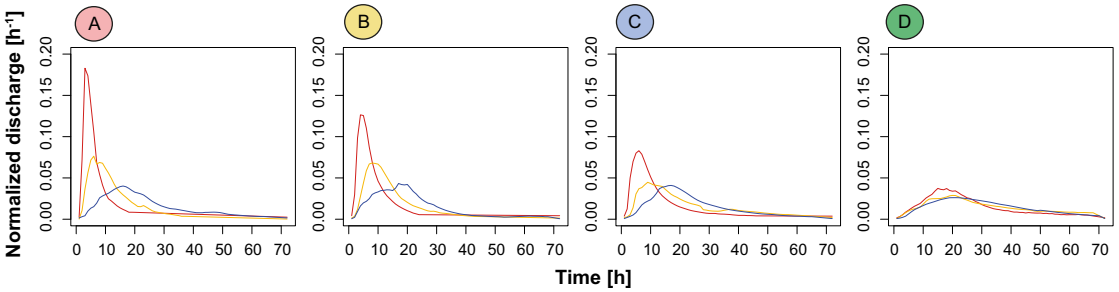



**Fig. 4.** Representative hydrograph shape sets per region and event type. The fast shapes are displayed in red, the intermediate shape in orange, and the slow shapes in blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 1**
Mean specific peak discharge ($Q_p$) [l/(s km$^2$)] and hydrograph volume ($V$) [m$^3$/km$^2$] per region (A to D) and event type (fast, intermediate, slow). The mean was computed from all observations within one event type.

| Event type | Region A | | Region B | | Region C | | Region D | |
|---|---|---|---|---|---|---|---|---|
| | $Q_p$ | $V$ | $Q_p$ | $V$ | $Q_p$ | $V$ | $Q_p$ | $V$ |
| Fast | 638 | 8607 | 365 | 7396 | 362 | 13,489 | 257 | 14,634 |
| Intermediate | 315 | 11,964 | 247 | 11,488 | 259 | 16,033 | 275 | 19,882 |
| Slow | 265 | 13,765 | 230 | 12,887 | 293 | 16,955 | 190 | 17,582 |

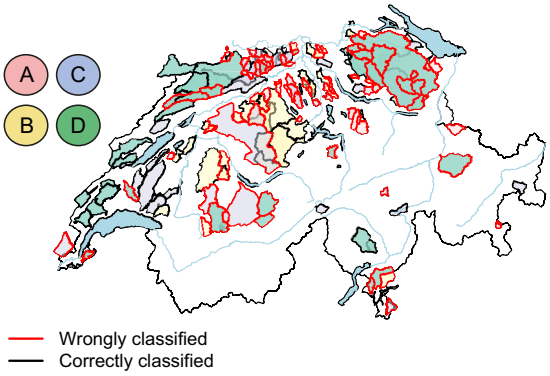



**Fig. 5.** 163 study catchments colored by their reactivity region membership: A: Quickly reactive, B: intermediately reactive, C: slowly reactive, and D: uniformly reactive. Catchments misclassified using the classification tree in the validation phase are indicated by red catchment borders, catchments correctly classified are indicated by black borders. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

generally higher for the fast SDHs than for the intermediate and slow SDHs. The variability of the relative error was larger towards negative values (i.e., underestimation of the regional estimates) and did not specifically depend on the hydrograph characteristics considered.

## 5. Discussion

### 5.1. Region assignment

Four regions with a similar flood behavior were identified. Catchments with a generally quick runoff response, independent of the event type, formed the quickly reactive region (Region A). Events occurring in this region are characterized by rather high peak discharges but rather low hydrograph volumes for all three event types. The catchments belonging to the quickly reactive region are small and mainly lie in the Swiss Plateau. They are characterized by a circular shape, impermeable rocks, a high network density, and low sunshine duration, which is related to their location in the Swiss Plateau that is often covered by fog in the winter time. All these characteristics contributed to a fast runoff reaction. A high network density allows for an efficient drainage after a precipitation event and is related to increases in the flood peaks (Ogden et al., 2011). Impermeable rocks prevent from rainfall infiltration and favor surface runoff, and a circular shape leads to the confluence of discharge from different parts of the catchment at the same time. Catchments in regions B and C with a generally intermediate or slow runoff response showed similar catchment characteristics. The events occurring in these catchments generally showed higher volumes but lower peak discharges compared to events occurring in the catchments of the quickly reactive region. These catchments typically lie in the Swiss Plateau or the Alps. Compared to the catchments in the quickly reactive region, they are more elongated, have more permeable rocks, and lower network densities which lead to a

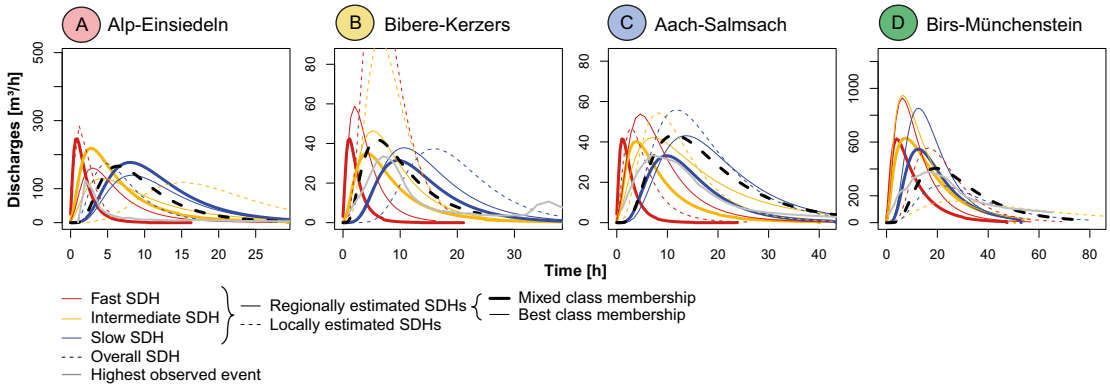



**Fig. 6.** Event-type specific SDHs for four example catchments for a return period of 100 years, one out of each reactivity region: Quickly reactive: A. Alp-Einsiedeln, intermediately reactive: B. Bibere–Kerzers, slowly reactive: C. Aach–Salmsach, uniformly reactive: D. Birs-Münchenstein. The regionally estimated event-type specific SDHs (bold lines) are plotted together with the locally estimated event-type specific SDHs (dashed lines). The catchment specific SDH not distinguishing between the event types is plotted in black. Regional estimates were derived based on the best region membership (thin lines) and using mixed class memberships (thick lines). The highest observed event in the individual catchments was added as a reference in grey.

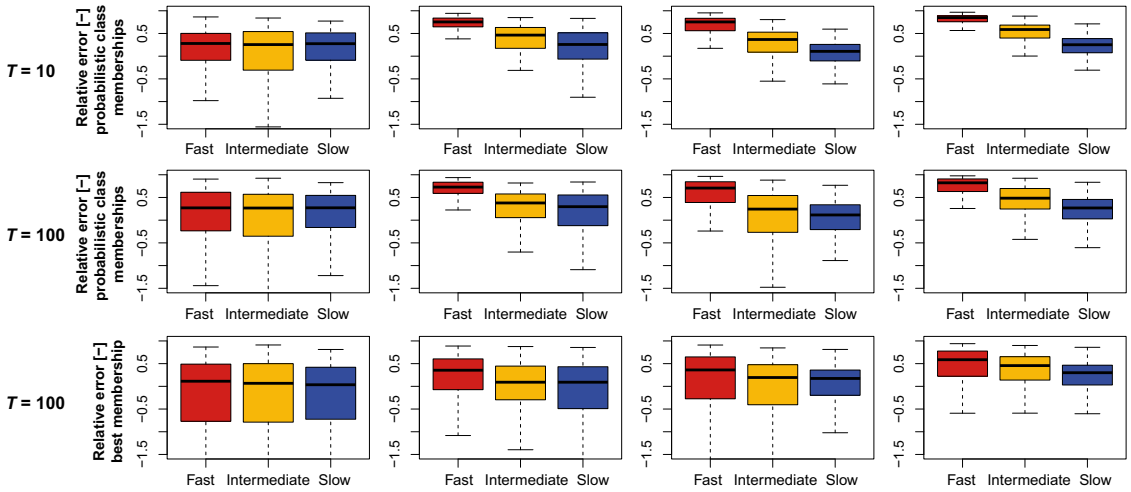



**Fig. 7.** Boxplots over all catchments summarizing the relative errors of regional event-type specific SDHs computed using three different configurations. The regional event-type specific SDHs shown in the upper panel were computed for a return period of 10 years using the probabilistic class memberships. The regional event-type specific SDHs shown in the middle panel were also computed using the probabilistic class memberships but for a return period of 100 years. The regional SDHs in the lower panel were computed for a return period of 100 years but applying the best class membership. The regional estimates were compared to local event-type specific SDHs for the respective return period (fast, intermediate, and slow) for four hydrograph characteristics: peak discharge, hydrograph volume, time to peak, and half-recession time. The whiskers extend to the lowest/highest data point which is still within 1.5 times the interquartile range. Outliers are not displayed.

dampening of the runoff events. Catchments with a uniform runoff reaction belong to region D, the uniformly reactive region. The events occurring in the catchments belonging to this region are generally characterized by slow rising and falling limbs, high volumes, and low peak discharges. The catchments belonging to this region are rather large and mainly lie in the Jura but also comprise a few large catchments in the Swiss Plateau. They are characterized by an elongated shape, karstic geology with permeable rock, and a low network density. These catchment properties lead to a damped runoff reaction that leads to similar flood events independent of the triggering mechanism and antecedent wetness. The mean floods of the three event types for the catchments in the four regions could be mainly explained by catchment area and network density.

We found that random forest was a suitable approach for establishing a classification rule allowing for the assignment of an ungauged catchment to one of the four reactivity regions even though the misclassification error seems rather high with 45%. However, not all misclassification errors were equally serious since we dealt with ordered classes. Assigning a catchment of Region A to Region B, for example, was less severe than assigning it to Region C because Regions A and B were more similar than Regions A and C. Most misclassification errors were related to placing a catchment into a neighboring region, more specifically to placing a catchment belonging to Region B into Region C or vice versa. These two classes were found to be difficult to distinguish using the classification rule. One could argue that these two regions could be combined as this would reduce the number of classes and therewith also the misclassification error. However, this combining would also increase the heterogeneity of flood characteristics of catchments within the combined class, which is not desirable for the SDH construction procedure. We therefore decided to keep the four classes and to benefit from the characteristics of the output of the ensemble-based random forest model. The ensemble-based method did not only provide a best class membership for a catchment but also probabilistic class memberships for each of the reactivity regions, which could be used in regional SDH construction. The weighting of estimates from different regions using the probabilistic region memberships helped to reduce the uncertainty related to misclassification and led to a reduction in the variability of the relative prediction error across catchments.

### 5.2. Regional event-type specific SDH sets

The four reactivity regions and the three event types were found to be useful for regional SDH construction since the events belonging to one event type within a region were found to be more similar than events belonging to another event type in the same region or the same event type in another region. Even though the individual data pools were not homogeneous in a statistical sense (Hosking and Wallis, 1997), they were still useful for regional flood frequency analysis. The different regional data pools did neither show strongly different forms nor intensities of dependence even though Grimaldi et al. (2016) suggested that basins with different times of concentration and different soil uses might show different dependence structures. This might be explained by the fact that the catchments under study were all located in the Swiss lowlands or Prealps which led to a certain similarity in land use and other catchment attributes. High alpine catchments, which were shown to have very low dependence between peak discharges and flood volumes due to a mix of flood types (Gaál et al., 2015), were not considered in the analysis.

The regional SDH construction approach relied on an index flood procedure that is applicable in ungauged catchments. The strongest explanatory variable for the index floods in the four regions was found to be catchment area, which is in line with findings by Blöschl and Sivapalan (1997) who found that there is a clear tendency of mean annual floods per area to decrease with catchment area. Potentially, this regional SDH construction approach can also be applied to catchments with only a few years of observations. There, the index floods could be derived from observed data and be combined with their regional growth curve. Furthermore, information could also be extended temporally (Merz and Blöschl, 2008) by including historical information (Wetter, 2017) which could potentially reduce the uncertainty in flood risk estimates (Kjeldsen et al., 2014).

The regional SDH construction approach was developed on a set of Swiss catchments but can be extended to other geographical regions for which a large enough dataset is available to delineate meaningful regions and to fit region specific generalized linear models for the prediction of index floods. However, a few assumptions would need to be verified and possibly adjusted. These comprise the number of flood reactivity regions, the fitting of the classification rule, the choice of catchment characteristics used in the classification model and the GLM for the prediction of the index flood, and the copula model used to

model the dependence between peak discharges and hydrograph volumes.

The regionally estimated event-type specific SDHs were compared to locally estimated event-type specific SDHs. This is not optimal since the locally estimated SDHs might not represent reliable estimates if they are estimated based on a very small flood sample. Still, the relative error of regionalized estimates compared to local estimates was quantified to lie around 50% on average across the 163 study catchments. The prediction errors were found to be generally higher for fast SDHs than for the intermediate and slow SDHs. This might be related to the fact that the slow events are more homogeneous across regions and event types than the fast events which generally show a higher variability.

## 6. Conclusions

The advantage of the event-type specific SDH construction procedure proposed here compared to a catchment-specific SDH construction procedure is that the three resulting SDHs provide us with a much better idea on the variability of potential design events. They depict a range of potential design flood outcomes showing the uncertainty related to different processes that can potentially occur within a catchment. This ensemble or set of design events can be used by engineers in hydraulic modeling or in a cost-benefit analyses when designing reservoir storage. The use of design flood ensembles in flood hazard mapping is advisable since flood peak attenuation varies with hydrograph magnitude and shape which could result in different water levels and flood extents. This ensemble-based design flood approach is a compromise between using a single best estimate design hydrograph and using a continuous simulation model for flood hazard mapping. Contrary to a best single estimate design hydrograph, it allows for the representation of process variability and does not give a result that is elusively precise. Still, compared to continuous models it does neither require the stochastic simulation of rainfall fields nor the calibration of a rainfall runoff model. It has the advantage that it is easily applicable in ungauged catchments and circumvents the simulation of rainfall and assumptions related to its transformation into runoff. However, setting up a regional index flood model might require as much data as setting up a regional continuous simulation model. The ensemble-based design hydrograph approach proposed in this study makes a step from a ”pure statistical approach” towards a ”process-based” method allowing for the representation of process variability in design flood estimation in ungauged catchments.

## Acknowledgments

We thank the Federal Office for the Environment (FOEN) for funding the project (contract 13.0028.KP / M285-0623) and for providing runoff measurement data. The data used in this study are available upon order from the FOEN. For the hydrological data of the federal stations, the order form under http://www.bafu.admin.ch/wasser/13462/13494/15076/index.html?lang=de can be used. The hydrological data of the cantonal stations can be ordered from the respective cantons. We thank the two reviewers for their constructive feedback.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.advwatres.2017.12.018

## References

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J.J., Capell, R., 2012. A comparison of similarity indices for catchment classification using a cross-regional dataset. Adv. Water Resour. 40, 11–22. http://dx.doi.org/10.1016/j.advwatres.2012.01.008.

Bitterli, T., Aviolat, P., Brändli, R., Christe, R., Fracheboud, S., Frey, D., George, M., Matousek, F., Tripet, J.P., 2007. Groundwater resources. Hydrological Atlas of Switzerland. Bern, p. 8.6.

Blöschl, G., Sivapalan, M., 1997. Process controls on regional flood frequency: coefficient of variation and basin scale. Water Resour. Res 33 (12), 2967–2980. http://dx.doi.org/10.1029/97WR00568.

Brunner, M.I., Seibert, J., Favre, A.-C., 2016. Bivariate return periods and their importance for flood peak and volume estimation. Wire's Water 3, 819–833. http://dx.doi.org/10.1002/wat2.1173.

Brunner, M.I., Viviroli, D., Furrer, R., Seibert, J., Favre, A.-C., 2017. Identification of flood reactivity regions via the functional clustering of hydrographs. Water Resour. Res. under review.

Brunner, M.I., Viviroli, D., Sikorska, A.E., Vannier, O., Favre, A.-C., Seibert, J., 2017. Flood type specific construction of synthetic design hydrographs. Water Resour. Res. 53. http://dx.doi.org/10.1002/2016WR019535.

Burn, D.H., Boorman, D.B., 1992. Catchment classification applied to the estimation of hydrological parameters at ungauged catchments. Technical Report. Institute of Hydrology, Wallingford, Oxfordshire.

Camezind-Wildi, R., 2005. Empfehlung Raumplanung und Naturgefahren. Technical Report. Bundesamt für Raumentwicklung, Bundesamt für Wasser und Geologie, Bundesamt für Umwelt, Wald und Landschaft, Bern.

Coles, S., 2001. An introduction to statistical modeling of extreme values. Springer, London.

Dalrymple, T., 1960. Flood-frequency analyses. Geological survey water supply paper 1543-A. pp. 80.

Deutsche Vereinigung für Wasserwirtschaft Abwasser und Abfall, 2012. Merkblatt DWA-M 552. Technical Report. DWA, Hennef, Germany.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrol. Process. 19, 507–515. http://dx.doi.org/10.1002/hyp.5675.

Eidgenössische Forschungsanstalt für Wald Schnee und Landschaft (WSL), 1999. Schweizerisches Landesforstinventar. Ergebnisse der Zwietaufnahme 1993-1995. BUWAL, Bern.

Frahm, G., Junker, M., Szimayer, A., 2003. Elliptical copulas: applicability and limitations. Stat. Prob. Lett. 63 (3), 275–286. http://dx.doi.org/10.1016/S0167-7152(03)00092-0.

Gaál, L., Szolgay, J., Kohnová, S., Hlavčová, K., Parajka, J., Viglione, A., Merz, R., Blöschl, G., 2015. Dependence between flood peaks and volumes - A case study on climate and hydrological controls. Hydrol. Sci. J. 60 (6), 968–984. http://dx.doi.org/10.1080/02626667.2014.951361.

Genest, C., Favre, A.-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Eng. 12 (4), 347–367. http://dx.doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347).

Gordon, A., 1999. Classification, 2nd ed. Chapman & Hall/CRC, Boca Raton.

Griffiths, G.A., 1989. A theoretically based wakeby distribution for annual flood series. Hydrol. Sci. J. 34 (3), 231–248. http://dx.doi.org/10.1080/02626668909491332.

Grimaldi, S., Petroselli, A., Salvadori, G., De Michele, C., 2016. Catchment compatibility via copulas: a non-parametric study of the dependence structures of hydrological responses. Adv. Water Resour. 90, 116–133. http://dx.doi.org/10.1016/j.advwatres.2016.02.003.

Grimaldi, S., Petroselli, A., Serinaldi, F., 2012. Design hydrograph estimation in small and ungauged watersheds: continuous simulation method versus event-based approach. Hydrol. Process. 26 (20), 3124–3134. http://dx.doi.org/10.1002/hyp.8384.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Harrell, F.E., 2015. Regression Modeling Strategies. With applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer, Cham.

Höllig, K., Hörner, J., 2013. Approximation and Modeling with B-Splines. Society for industrial and applied mathematics, Philadelphia.

Hosking, J., 1986. The Wakeby distribution. Technical Report. IBM T.J. Watson Research Center, Wallingford UK.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511529443.

Houghton, J.C., 1978. Birth of a parent: the wakeby distribution for modeling flood flows. Water Resour. Res. 14 (6), 1105–1109. http://dx.doi.org/10.1029/WR014i006p01105.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning. With Applications in R. Springer, New York. http://dx.doi.org/10.1007/978-1-4614-7138-7.

Kjeldsen, T.R., Macdonald, N., Lang, M., Mediero, L., Albuquerque, T., Bogdanowicz, E., Brazdil, R., Castellarin, A., David, V., Fleig, A., Gül, G.O., Kriauciuniene, J., Kohnova, S., Merz, B., Nicholson, O., Roald, L.A., Salinas, J.L., Sarauskiene, D., Sraj, M., Strupczewski, W., Szolgay, J., Toumazis, A., Vanneuville, W., Veijalainen, N., Wilson, D., 2014. Documentary evidence of past floods in europe and their utility in flood frequency estimation. J. Hydrol. 517, 963–973. http://dx.doi.org/10.1016/j.jhydrol.2014.06.038.

Lang, M., Ouarda, T., Bobée, B., 1999. Towards operational guidelines for over-threshold modeling. J. Hydrol. 225, 103–117.

Lettenmaier, D.P., Wallis, J.R., Wood, E.F., 1987. Effect of regional heterogeneity on flood frequency estimation. Water Resour. Res. 23 (2), 313–323. http://dx.doi.org/10.1029/WR023i002p00313.

Madsen, H., Rasmussen, P.F., Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. at-site modeling. Water Resour. Res. 33 (4), 747–757.

Merz, R., Blöschl, G., 2003. A process typology of regional floods. Water Resour. Res. 39 (12), 1340. http://dx.doi.org/10.1029/2002WR001952.

Merz, R., Blöschl, G., 2008. Flood frequency hydrology: 1. temporal, spatial, and causal expansion of information. Water Resour. Res. 44. http://dx.doi.org/10.1029/2007WR006744.

MeteoSwiss, 2013. Documentation of MeteoSwiss grid-data products: Daily precipitation (final analysis): RhiresD. Technical Report. MeteoSwiss.

Myers, R.H., Montgomery, D.C., Vining, G.G., Robinson, T.J., 2010. Generalized Linear Models. Vol. 4 John Wiley & Sons, Inc, Hoboken, New Jersey.

Nadarajah, S., 2007. Probability models for unit hydrograph derivation. J. Hydrol. 344, 185–189. http://dx.doi.org/10.1016/j.jhydrol.2007.07.004.

Ogden, F.L., Raj Pradhan, N., Downer, C.W., Zahner, J.A., 2011. Relative importance of impervious area, drainage density, width function, and subsurface storm drainage on flood runoff from an urbanized catchment. Water Resour. Res. 47 (12). http://dx.doi.org/10.1029/2011WR010550.

Oudin, L., Kay, A., Andréassian, V., Perrin, C., 2010. Are seemingly physically similar catchments truly hydrologically similar? Water Resour. Res. 46, W11558. http://dx.doi.org/10.1029/2009WR008887.

Pilgrim, D.H., 1986. Bridging the gap between flood research and design practice. Water Resour Res 22 (9), 165–176.

Rai, R.K., Sarkar, S., Singh, V.P., 2009. Evaluation of the adequacy of statistical distribution functions for deriving unit hydrograph. Water Resour. Manag. 23, 899–929. http://dx.doi.org/10.1007/s11269-008-9306-0.

Requena, A.I., Chebana, F., Mediero, L., 2016. A complete procedure for multivariate index-flood model application. J. Hydrol. 535, 559–580. http://dx.doi.org/10.1016/j.jhydrol.2016.02.004.

Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., Merz, R., Blöschl, G., 2012. Runoff models and flood frequency statistics for design flood estimation in austria - Do they tell a consistent story? J. Hydrol. 456, 30–43. http://dx.doi.org/10.1016/j.jhydrol.2012.05.068.

Schmid, F., Schmidt, R., 2007. Multivariate conditional versions of Spearman's rho and related measures of tail dependence. J. Multivar. Anal. 98 (6), 1123–1140. http://dx.doi.org/10.1016/j.jmva.2006.05.005.

Sikorska, A.E., Viviroli, D., Seibert, J., 2015. Flood type classification in mountainous catchments using crisp and fuzzy decision trees. Water Resour. Res. 51 (10), 7959–7976. http://dx.doi.org/10.1002/2015WR017326.

Wetter, O., 2017. The potential of historical hydrology in switzerland. Hydrol. Earth Syst. Sci. Discuss. 27–34. http://dx.doi.org/10.5194/hess-2017-410.

Yue, S., Ouarda, T., Bobée, B., Legendre, P., Bruneau, P., 2002. Approach for describing statistical properties of flood hydrograph. J. Hydrol. Eng. 7 (2), 147–153. http://dx.doi.org/10.1061/(ASCE)1084-0699(2002)7:2(147).